Method Development in Crystallization for Femtosecond Nanocrystallography

by

Christopher Kupitz

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2014 by the
Graduate Supervisory Committee:

Petra Fromme, Co-Chair
John C. Spence, Co-Chair
Kevin Redding
Alexandra Ros

ARIZONA STATE UNIVERSITY

August 2014

ABSTRACT

Membrane proteins are a vital part of cellular structure. They are directly involved in many important cellular functions, such as uptake, signaling, respiration, and photosynthesis, among others. Despite their importance, however, less than 500 unique membrane protein structures have been determined to date. This is due to several difficulties with macromolecular crystallography, primarily the difficulty of growing large, well-ordered protein crystals. Since the first proof of concept for femtosecond nanocrystallography showing that diffraction patterns can be collected on extremely small crystals, thus negating the need to grow larger crystals, there have been many exciting advancements in the field. The technique has been proven to show high spatial resolution, thus making it a viable method for structural biology. However, due to the ultrafast nature of the technique, which allows for a lack of radiation damage in imaging, even more interesting experiments are possible, and the first temporal and spatial images of an undamaged structure could be acquired. This concept was denoted as time-resolved femtosecond nanocrystallography.

This dissertation presents on the first time-resolved data set of Photosystem II where structural changes can actually be seen without radiation damage. In order to accomplish this, new crystallization techniques had to be developed so that enough crystals could be made for the liquid jet to deliver a fully hydrated stream of crystals to the high-powered X-ray source. These changes are still in the preliminary stages due to the slightly lower resolution data obtained, but they are still a promising show of the power of this new technique. With further optimization of crystal growth methods and quality, injection technique, and continued development of data analysis software, it is

only a matter of time before the ability to make movies of molecules in motion from X-ray diffraction snapshots in time exists. The work presented here is the first step in that process.

I dedicate this thesis to my wife, Emily A. Kupitz, and my son Benjamin J. Kupitz, without whom this would never have been possible. Thank you both so very much for sharing my life with me.

ACKNOWLEDGEMENTS

I would first like to acknowledge my supervisor, Dr. Petra Fromme, for a fantastic four years supporting me and teaching me in a very complicated field. She exhibits an enthusiasm that makes you want to come into work and an environment that makes you enjoy it when you are there. She has taught me what it means to have enthusiasm for the project you are working for that I will carry with me forever.

I am also very grateful to Dr. John C. H. Spence, for all of his support and advice throughout this work, and especially his encouragement whenever things were going tough. It was exactly what I needed in order to get back up and try again when things weren't working the way we wanted them to. He has made an invaluable impression on me as a scientist and I was honored to work with him.

To my committee, Dr. Alexandra Ros and Dr. Kevin Redding, I would extend my heartfelt appreciation for all of their advice whenever I had strange ideas or experiments and they helped me to make them better and make them work. Thank you for teaching me so many things during this time.

A great deal of the success of this project goes to Shibom Basu. He was the main data analyst who helped to make heads or tails out of the data that we generated together, and with his help we made improvements to the crystals based upon that data. Thank you very much for all your time and discussions with me about this work.

I owe a huge debt to Dr. Ingo Grotjohann for all of his help through the difficult period before a beamtime in regards to cell growth, protein preparation, and help with discussing chemistry techniques in general and crystallization techniques in particular. Thank you to Dr. Raimund Fromme for all of his help in freezing crystals, collecting data

sets, and helping to analyze the extensive amount of data that has been collected. Thank you to my fellow graduate students, Chelsie E. Conrad, Shatabdi Roy-Chowdhury, and Jesse Coe, who have helped me perform countless experiments and preps and who have had many discussions with me about the science behind what we are doing. I would like to thank Dr. James Zook and Dr. Mark Hunter for all of their help in teaching me theory and how to begin my experiments when I first started at ASU and was lost. I couldn't have asked for better friends to guide me through this.

This project was a monumental effort that requires the hard work of many people, especially in the Physics Department at Arizona State University. Dr. Uwe Weierstall was always being willing to discuss the project with me, and he had constant patience in explaining the injector so that I looked like I knew what I was doing. Thank you to both Dingjie Wang, and Daniel James for their tireless efforts to keep the experiments for this project going and for always being willing to lend a hand to whatever new idea it was I wanted to discuss or try. I also need to mention Dr. Nadia Zatsepin who continuously had to explain all of the theories behind the data analysis to me repeatedly so that I would understand what they were doing, and who is always willing to lend an ear when I need to discuss some aspect of the project. Thank you to you all; you are a huge part of the reason this project has been a success so far.

This project was part of a large international collaboration that was formed to work on femtosecond X-ray protein nanocrystallography. All of the authors on the 2014 Nature Letter deserve my thanks. Specifically I would like to mention Henry Chapman, Mike Bogan, Ilme Schlickting, Robert Shoeman, Anton Barty, Tom White, Andy Aquila, Raymond Sierra, Hartawan Laksmono and Francesco Stellato. I would like to thank Mike

Bogan for allowing us to take over his laboratory at the Pulse Institute at Stanford frequently when we were running experiments, as well as for very interesting discussions on various topics in science. Thank you to Robert Shoeman for many helpful ideas, always advocating our work, always having good comments and fast solutions to frequently difficult problems, and most of all for teaching me many of the facets of sample loading and injection that I found fascinating. I would like to thank Francesco Stellato, Hartawan Laksmono, and Raymond Sierra, for much help and discussions regarding sample preparation and characterization. I would like to thank Tom White who played a major part in writing the software that we use for data analysis, as well as for always being willing to explain whatever it was that I didn't understand during a beamtime. I would like to thank Andrew Aquila for all of his help in "piloting" an experiment, as well as for teaching me it really isn't as easy as it looks. I would like to thank Ilme Schlichting, for all of her help in every step of beamtimes and for her discussions and willingness to listen to whatever the problem may be and help with her knowledge to solve it. It was always enjoyable having a discussion with her. I would like to thank Henry Chapman for initially bringing together this large and disparate collaboration which has made this work possible, and for continuously providing necessary insight and creative ideas to solve any problems that come up. I have found it to be a joy to work with you all during this time.

Experiments were carried out at the Linac Coherent Light Source (LCLS), a national user facility operated by Stanford University on behalf of the US Department of Energy (DOE), Office of Basic Energy Sciences (OBES).

This work was supported by the following agencies: the Center for Bio-Inspired Solar Fuel Production, an Energy Frontier Research Center funded by the DOE, Office of Basic Energy Sciences (award DE-SC0001016), the National Institute of Health (award 1R01GM095583), the US National Science Foundation (award MCB-1021557 and MCB-1120997), the DFG Clusters of Excellence "Inflammation at Interfaces" (EXC 306) and the "Center for Ultrafast Imaging"; the Deutsche Forschungsgemeinschaft (DFG); the Max Planck Society, the Atomic, Molecular and Optical Sciences Program; Chemical Sciences Geosciences and Biosciences Division, DOE OBES (M.J.B) and the SLAC LDRD program (M.J.B, H.L); the US DOE through Lawrence Livermore National Laboratory under the contract DE-AC52-07NA27344 and supported by the UCOP Lab Fee Program (award no. 118036) and the LLNL LDRD program (12-ERD-031); the Hamburg Ministry of Science and Research and Joachim Herz Stiftung as part of the Hamburg Initiative for Excellence in Research. I also want to thank the National Science Foundation for providing funding for the publication of this work through the BioFEL Science Technology Center (award 1231306).

I would like to specially thank my family, my parents and brother for lending an ear when it was needed even though they were probably bored to tears. They are the people who started me out on this path and supported me.

The biggest thanks of all goes to my wife and son who have dealt with everything I have had to and who keep offering me their support and love. They are the best things that have ever happened to me.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Chapter 1

Introduction

Introduction to Macromolecular X-ray Crystallography

Principles of X-ray Crystallography and Accomplishments

X-ray crystallography is currently the predominant technique for solving protein structure, and has solved 90,000 structures in the Protein Data Bank, compared to 10,500 structures from nuclear magnetic resonance and 1,000 structures of electron microscopy. The basic premise of X-ray crystallography is a crystal's ability to coherently scatter electromagnetic radiation. Unfortunately, proteins are composed of mostly H, C, N, and O atoms. These atoms have a relatively low atomic number, so the scattering intensity is weak(Hunter et al., 2011). One of the major bottlenecks in structure determination by X-ray crystallography is the growth of large, well-ordered crystals (Bill et al., 2011) needed to increase scattering intensities. Membrane proteins are particularly problematic because they are of amphiphilic nature with large parts of the hydrophobic protein being embedded into a membrane. These proteins are solubilized from the membrane and crystallized whether in the form of protein-detergent micelles or crystallized in lipidic phases (Cherezov et al., 2001; Hunter and Fromme, 2011). Due to these difficulties, of the >100,000 protein structures in the PDB (Protein Data Bank), only about 500 unique membrane protein structures have been solved, despite membrane proteins comprising approximately 30% of all unique proteins.

Photosystem I (PSI) from *Thermosynechococcus elongatus* is an excellent example of why so few membrane protein structures have been solved. For PSI, which is a huge membrane protein complex of 1 million MW, the first report of a shower of micro-sized crystals was reported in 1988 (Witt et al., 1988). From there it was another five years before a first low resolution structure was solved to 6Å in 1993 (Krauss et al., 1993; Witt et al., 1994). In 1996 the resolution was improved to 4Å (Krauss et al., 1996), based on a native data set from a large single crystal grown under microgravity. A further improved structure was published in 1999 (Klukas et al., 1999). Finally in 2001, the structure of PSI was solved to 2.5Å using cryogenically frozen crystals (Jordan et al., 2001). Thus it took thirteen years from the first discovery of very small crystals until a high resolution structure became available (Fromme and Mathis, 2004). This was a huge accomplishment taking many researchers over a decade to finally complete, and this is only one example of both the difficulties and successes associated with macromolecular crystallography.

The reasons why macromolecular crystallography can be so difficult can generally be divided into three broad, separate areas of challenges: 1. expression, purification and isolation of protein in mono-disperse, functional intact form, 2. the ordered growth of large crystals, and 3. the problem of radiation damage that leads to photodamage of the protein and the crystals during exposure to X-rays. The discussion will now focus on the second two points, which are directly related to X-ray structure determination.

The growth and nucleation of protein crystals depends on multiple factors, but one of the primary factors is the supersaturation of the protein solution. To understand

this, one must first understand the phase diagram. The phase diagram is a schematic, typically of precipitant concentration vs. protein concentration. On this graph there are normally four different zones: the unsaturated, where everything is stable, the metastable, where crystal growth occurs, the nucleation where the crystals originally start growing, and the precipitation, where proteins come together in a non-ordered packing. For an example of this, see Fig. 2.2 in the second chapter.

At high supersaturation, when the solution reaches the nucleation zone, nulei are larger than the critical radius, so crystal growth is observed. However, it is difficult to grow large well-ordered crystals, as nucleation and growth of existing crystals compete for the proteins in solution. Therefore, ideally nucleation is induced in the nucleation zone followed by growth of crystals in the metastable zone. However, if the supersaturation has only reached the nucleation zone, nuclei are not stable so no induction of nucleation occurs, while pre-formed crystals can grow. If the supersaturation reaches the border between nucleation and precipitation zone, crystals growth is extremely fast, leading to highly disordered crystals. When the supersaturation is too high the protein may just form amorphous precipitate.

Even if a large crystal can be grown, they often have a high degree of internal disorder. The degree of disorder is determined from the broadening of the Bragg peaks on the diffraction pattern. The classical explanation refers to disorder in crystals as formation of "mosaic blocks," which explains why the disorder is expressed by a term called "mosaicity" (Woolfson, 1997). Recent literature contains reports where actual images were taken from large crystals that consist of smaller crystalline blocks (Nederlof et al., 2013), and each block can have a small misalignment from the blocks next to it,

leading to mosaicity. This will have an effect of broadening the reflections during data acquisition, which can make it very difficult to determine the amplitudes of the structure factors from the diffraction patterns.

The X-ray Damage Problem in Macrocrystallography

The largest problem facing macrocrystallography is the radiation damage which occurs because X-rays are ionizing radiation. The dose-damage ratio finally limits the size of a crystal from which X-ray data can be collected. The literature discriminates between two major causes of radiation damage in protein crystallography: local and global damage (Holton, 2009). Local damage is a form of damage that affects very specific sites such as metal atoms or clusters, or certain amino acids, especially amino acids containing sulfur (Cys and Met), which are prone to X-ray damage (Holton, 2009). Global damage, on the other hand, is not associated with any specific element of the protein, but is instead an overall effect of damage over time to the entire crystal. On a molecular level, radicals are formed leading to, among other problems, photoreduction of metals and breakage of chemical bonds leading to destruction of the protein. This manifests in multiple ways and includes increased unit cell size, increased internal disorder of the crystals, amongst many others, but is dependent upon the conditions of the experiment (Meents et al., 2010). Howells et al. has defined a maximal X-ray dose for diffraction of a biomolecule to be 10 MGy per 1 Å of resolution generally (Howells et al., 2009). Many of the most interesting proteins contain cofactors that are even more strongly affected by X-ray damage and decay significantly before this damage threshold. For example, in Photosystem II (PSII) the Mn cluster, which is the oxygen evolving

center, is very strongly affected by photoreduction of the metals, which makes it very difficult to solve the native undamaged structure of PSII.

Radiation damage is primarily caused by inelastic events, namely Compton scattering and the photo-electric effect. These effects can cause the ejection of electrons by photons, causing a cascade of photoelectric events. This initial event and the following cascade can cause extensive damage, including chemical bond breakage, free radical generation, and oxidation-reduction reactions. All of this damage is accumulated at the microscopic level, which leads to further long-range disorder, or global damage (Meents et al., 2010).

Specific damage, on the other hand, happens primarily to heavier atoms in the protein, which includes metal cofactors and amino acids that contain sulfur. Generally affected sites are disulfide bonds, but also decarboxylation; and loss of hydroxyl groups are observed. Metals and cofactors tend to be more prone to radical formation and thus also accrue a higher portion of damage from X-ray exposure.

In an effort to mitigate the radiation damage and thus increase the ability to collect better data it was found that cooling the sample led to a slowing of the secondary radiation damage (Haas and Rossmann, 1970). This led to the development of cryogenic cooling of the protein crystals as the routine way to collect X-ray diffraction data on protein crystals (Hope, 1988). The result of this cryogenic crystallography is the decreased diffusion rate of radicals, which limits the amount of damage and decreases the speed at which the long range order of the crystal decays in the X-ray beam. Unfortunately, there is still an absolute limit on the amount of radiation damage that can

be suppressed and therefore large crystals are still needed in order to withstand the radiation dose(Yano et al., 2005; Howells et al., 2009; Meents et al., 2010).

Cryogenic cooling of protein crystals is now routinely done for X-ray data collection; however it has several major limitations. First is that most protein crystals cannot be directly frozen from their mother liquor; they must be treated with some form of cryo-protectant i.e. glycerol, sucrose, etc., in order to avoid the formation of ice. The presence of cryo-protectant will add significantly to the background, but the addition of the cryo-protectant can also change the chemical composition of the solution the crystals are in, and thus may dissolve or lower the order of the crystal. Another major problem is that proteins may become locked in non-native conformations, thereby giving a structure that is not actually biologically relevant. This would render the experiment less than optimal as the protein structure is solved with the goal to understand the protein function. Unfortunately these problems are unavoidable and a lot of time and research effort has been dedicated to eradicating these issues.

Diffraction before Destruction – Damage Free Imaging?

A breakthrough in thinking about damage reduction occurred in the year 2000 with the publishing of the paper by J Hajdu and R Neutze, where they simulated the damage caused to a single lysozyme molecule in vacuum exposed to an ultra-short femtosecond X-ray pulse with X-ray photon flux that could be reached by a X-ray Free Electron Laser with a peak power that would be 10^{12} higher than a synchrotron. This study proposed that in theory it should be possible to collect damage free X-ray diffraction data using pulses of X-rays (Neutze et al., 2000). This simulation showed a lysozyme molecule, in the gas phase, interacting with a very intense X-ray pulse from an

XFEL. This led to a large positive charge on the lysozyme which in turn led to a Coulomb explosion. The interesting thing about this simulation is that the resulting explosion took place on a 5-10 femtosecond timescale, seen in Fig. 1.1. This was then interpreted that if the X-ray pulse were to be shorter than 5 femtoseconds the diffraction pattern could be collected before damage occurs. This led to the proposed

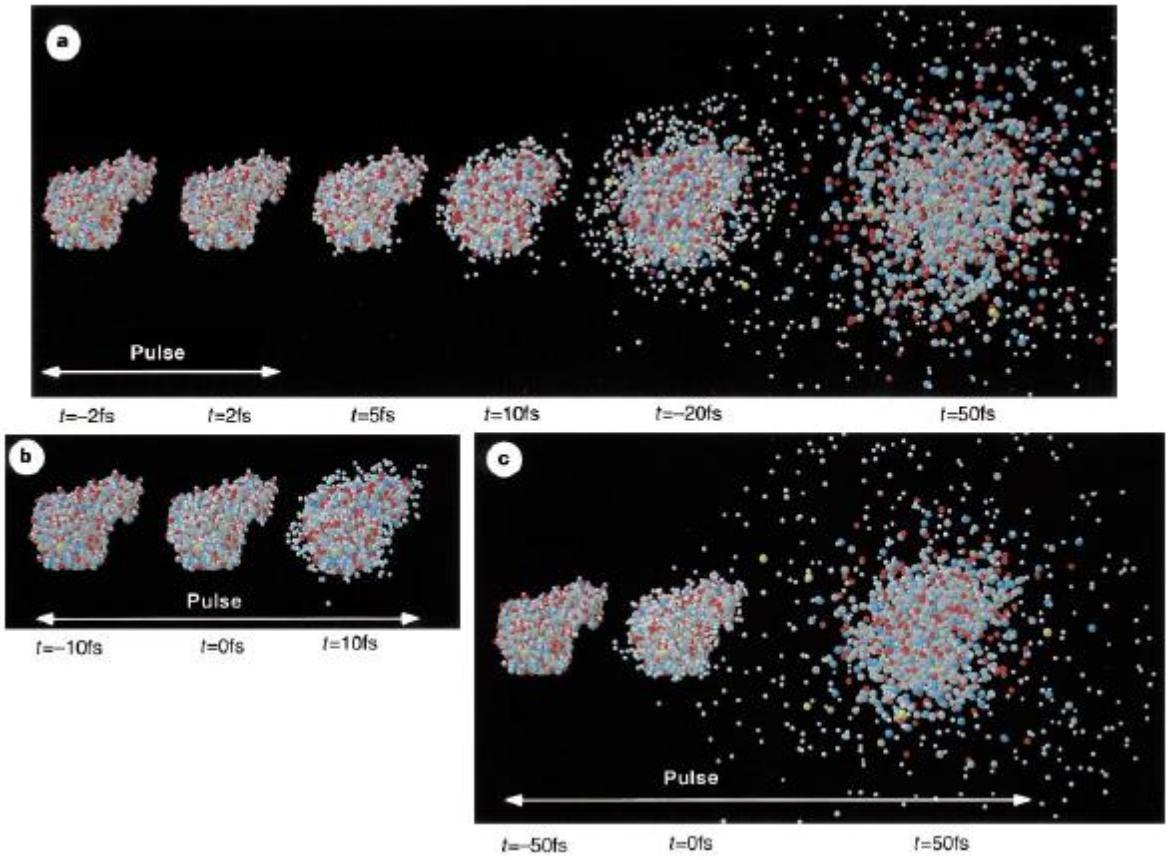


Figure 1.1. ***Coulomb Explosion Simulation****: This is an explosion of a T4 lysozyme (white, H; grey, C; blue, N; red, O; yellow, S) that is induced by radiation damage. The X-ray intensity was 3x10^{12} (12 KeV) photons/ 100-nm diameter spot (3.8x10^{6} photons/ Å^{2}). (a) Protein exposed to an X-ray pulse at FWHM of 2 fs. The disintegration of the molecule is followed in time. The atomic positions before and after (first two time points) remain virtually unchanged at this pulse length, due to the inertial delay in the explosion Rnucl = 3%, Relec = 11% (b) Lysozyme shown the same number of photons as (a), but FWHM (Full width half max) pulse is 10 fs. Structure shown at the beginning, middle and end of the pulse Rnucl = 7%, Relec = 26%. (c) Lysozme shown a pulse with FWHM of 50 fs. Rnucl = 26%, Relec = 30%. Figure taken and caption adapted from(Neutze et al., 2000).*

"diffract-before-destroy" principle, a theory in which X-ray radiation damage is avoided by ultra-short X-ray pulse duration. At the time when this work was originally published, however, there existed no X-ray source at which it could be tested, so there was no way to prove the theory in 2000.

Free Electron Lasers Open New Technical Advantages for Structural Biology

A Brief History of X-rays and X-ray Sources

X-rays were first reported by a German physicist, Rector Roentgen, in 1895 (Hauptman, 1991). His discovery came from one of the major experimental traditions in the latter half of the nineteenth century, which was to study the discharge of electricity in gases. Within a month of his discovery, X-rays were in use all over the world as a method for doctors to identify broken bones and find bullets. In 1899 Max von Laue proposed an idea that the regularly spaced atoms in a crystal would make a good diffraction grating for X-rays (Friedrich et al., 1912), which would diffract at ultra-high frequencies . At the time it was widely believed that X-rays were inherently electromagnetic in nature, and therefore the set planes between atoms imposed structure upon the X-ray pulse. This led Bragg to propose the following equation that defines under which condition a set of planes will lead to diffraction (known now the Bragg Law), see Equation 1,

Equation 1: $$n\lambda = 2dsin\theta$$

where d is the distance between planes and θ is the angle of reflection. These original experiments were performed using X-ray tubes as an X-ray source, which is an enclosed anode and cathode set-up where the cathode was heated via electrical current to produce

electrons from a tungsten filament that were then accelerated toward the anode where X-rays were emitted. Initially various minerals were studied and the data was collected on photographic plates. In 1912, Lawrence Bragg derived the formulas by which structures could be solved from X-ray diffraction patterns. By 1913, X-ray crystal analysis had become a standard technique, merely a year after its discovery. In 1934 the first diffraction from a small protein crystal, pepsin, was obtained (Bernal and Crowfoot, 1934; Kendrew et al., 1958), with the first actual structure from myoglobin solved in 1958 (Bodo et al., 1959). The first enzyme, lysozyme, was then solved in 1965 (Blake et al., 1965; Johnson and Phillips, 1965).

Synchrotrons were originally developed for high powered particle physics (Rowe and Mills, 1973; Kohoutova et al., 2009). They are essentially a type of cyclic particle accelerator. Synchrotrons, also called storage rings, use a magnetic field to guide a particle beam of increasing kinetic energy. In order to generate X-rays, a high powered particle beam, in this case an electron beam, is guided through bending magnets, undulators, or wigglers, whereby X-rays are emitted, which converts the electron energy into photons (Bilderback, 2005), thus giving off synchrotron radiation in the form of X-rays. Even these first generation sources were considerably more powerful, approximately 10^{14} more peak brilliance(Altarelli, 2010), than the X-ray tubes that had been previously used. Further advancements to synchrotrons have been made during the second and third generation by increasing the storage ring size, adding an additional storage ring to increase power, adding a linear accelerator, and many other ways. The increase in power with each additional advancement is both a blessing and a curse. With an increase in flux there is an increase in intensity of the diffraction, which allows for the

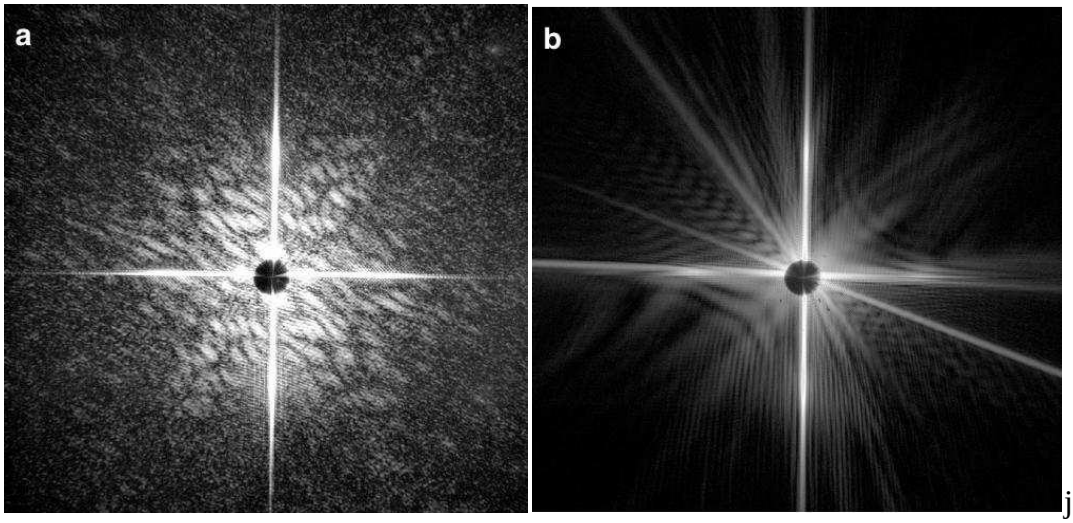
use of smaller crystals. However, it also increases the radiation damage suffered by the crystal, which calls for data collected on multiple small crystals or data collection on ever larger crystals that can be shifted during data collection to limit the increase in X-ray damage.

The Advent of X-ray Free Electron Lasers

As limits were approached with what synchrotrons could reach in terms of improvements, a new technology came onto the scene. This was the X-ray free electron laser (FEL), the first of which was built at the Deutsches Elektronen-Synchrotron (DESY) in Hamburg and was called the Free electron LASer in Hamburg (FLASH). FLASH started operation in 2005 with a short X-ray pulse duration of 10-50 femtoseconds (fs), and a wavelength of between 47 and 470 Å. This makes it a soft X-ray facility unable of solving structures to a high resolution, but it provided the very important first low resolution X-ray diffraction data from X-ray diffraction of cells, viruses and many other samples using an X-ray FEL(Chapman et al., 2006; Bostedt et al., 2012; Zastrau et al., 2014). This was the first time the diffract-before-destroy principle was tested. The first sample for this test was a picture of a man and a sun etched into a silicon nitride window. With this sample they aimed to determine whether a diffraction pattern collected of the sample by a single fs X-ray pulse at FLASH could actually produce an undamaged image of the design. (Chapman et al., 2006). An example of the diffraction pattern and a single particle reconstructed image can be seen in Figs. 1.2 and 1.3 respectively. Unfortunately, due to the limited resolution available at FLASH, the

question whether the damage-free reconstruction would extend to higher resolution was still open.

The Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory was the second X-ray FEL worldwide; it started operation in April 2009. LCLS was the world's first high-energy FEL, which provided photon energies of up to 2keV and thereby reached a wavelength of 6.9 Å, which was relevant to biological problems. The first experiments at LCLS (Emma, 2010) were performed at the Atomic and Molecular Optics instrumental station (beamline AMO) at the end of 2009. The



j

Figure 1.2. ***Coherent X-ray Diffraction Patterns from FLASH****. (a) Coherent diffraction pattern recorded for a single (4 ± 2) × 10^{14} W/cm^2, 25 ± 5 fs pulse, and (b) for the subsequent pulse of similar intensity and duration, 20 s later, showing diffraction from the damage caused by the pulse that formed (a). The intensity is shown on a logarithmic grey scale with black denoting 1 photon/pixel and white denoting 2000 photons/pixel for (a) and 50,000 photons/pixel for (b). The entire patterns are shown as detected by the CCD, and extend to a diffraction angle of 15◦ at the midpoint of the edges (corresponding to a momentum transfer of 8.1 μm-1). Image and caption taken and modified from(Chapman et al., 2006).*

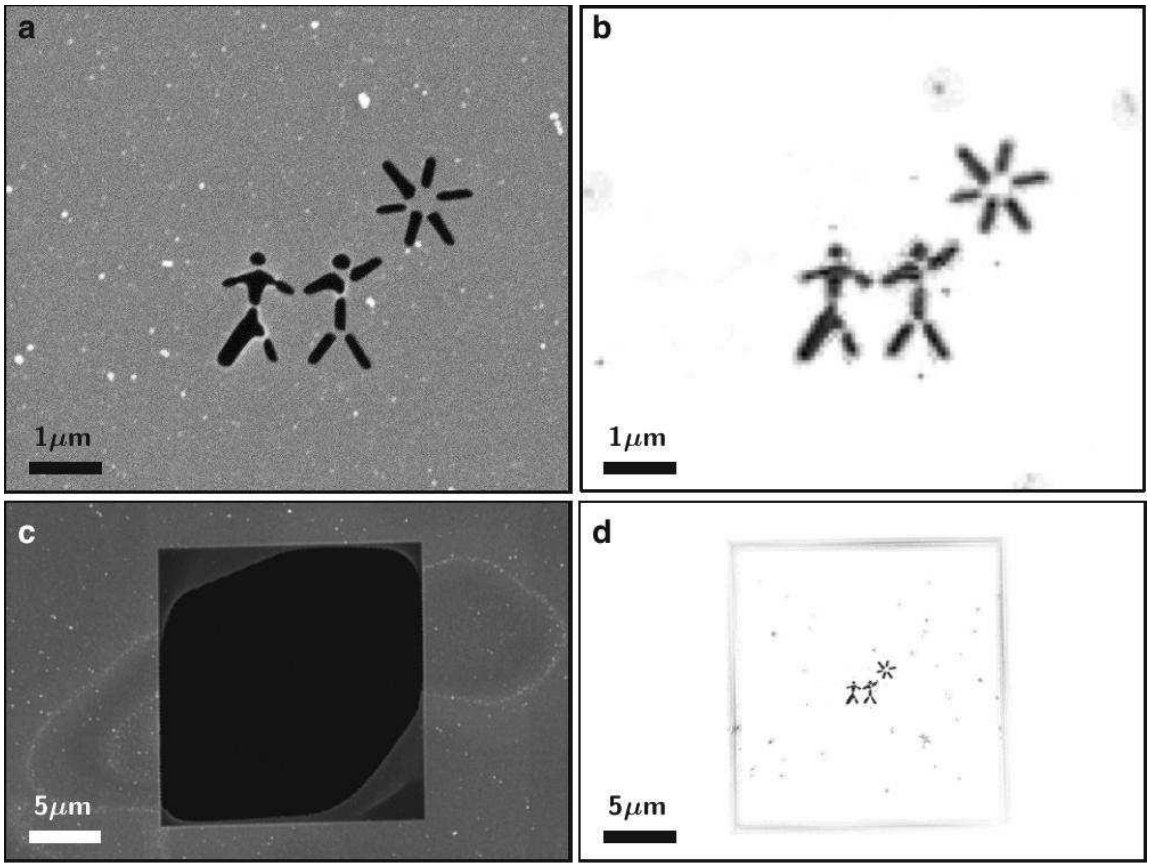


Figure 1.3. ***Reconstruction of a Sample Image without Damage****: (a) Scanning electron microscope (SEM) image of the sample (two men and a sun etched into a silicon nitrate window) before exposure to the FEL beam. The 20 nm thick sample was held in a square supporting window that is 20 μm wide. (b) and (d) Image reconstructed, from the ultrafast coherent diffraction pattern of Fig. 1.2 (a), by phase retrieval and inversion using the Shrinkwrap algorithm(Marchesini et al., 2003). The squared modulus of the retrieved complex image is displayed on a linear grey scale from zero scattered photons/pixel (white) to 1.5 × 106 scattered photons/pixel (black). Pixel size in the reconstruction = 62 nm in (b), corresponding to the half period of the finest spatial frequency that can be recorded on our camera at 32 nm wavelength. The retrieved image clearly shows the silicon window edge (in d), the FIB pattern, and dirt particles. (c) SEM image of the test sample after the exposures to the FEL beam, showing the square 20-μm window and some remaining silicon nitride, as well as visible damage to the silicon support caused by the non-circular beam. The scale bar for (a) and (b) is 1 μm and the scale bar for (c) and (d) is 5 μm. Image and caption taken from (Chapman et al., 2006)*

initial wavelength of the X-ray laser was only to 6.9 Å with a repetition rate of 30 Hz and a flux of 10^{12}-10^{13} photons/pulse. For a comparison between X-ray sources including

FELs, see Fig. 1.4. The first experiments performed on protein crystals were done at AMO, showing damage-free data collection on the large membrane protein complex Photosystem I (Chapman 2011). The fact that diffraction data could be collected from nanocrystals of the extremely large and difficult membrane protein complex showed proof of principle that the diffract-before destroy can be applied to protein crystals (Barty et al.) and that the new method of serial femtosecond crystallography is a viable technique that can solve structures from difficult-to-crystallize proteins. The most interesting experimental station at LCLS for scientist working on structural biology is the Coherent X-ray Imaging (CXI) beamline (Boutet and Williams, 2010). This beamline became available for users in February 2011. The opening of CXI went hand in hand with the increase of the energy of LCLS to > 10keV, which allowed for the first time data collection at atomic resolution (Boutet et al.). Furthermore, the repetition rate of the FEL was increased to 120 Hz repetition rate with the maximal flux being maintained at 10^{12} photons/pulse. This advancement brought the possibility for high resolution data collection from an X-ray FEL for the first time, thus creating a new field, serial femtosecond nanocrystallography (SFX).

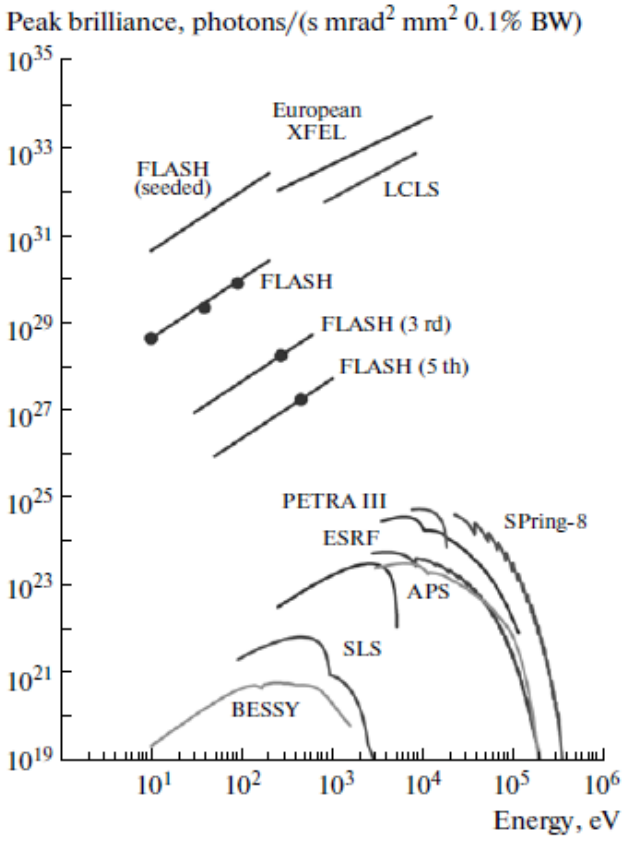


Figure 1.4. ***Peak brilliance plotted against photon energy*** *of both XFELs (planned and built) and third generation synchrotron light sources. Dots denote actual measured values, various harmonics for FLASH undulator is also shown. Figure and caption taken and modified (Altarelli, 2010).*

Current/Future FELs under Construction

Since the first successful experiments have been performed at an X-ray FEL, many FELs are in the planning or construction stage. One has already been completed, three are under construction, and several more are just in the planning stage. For example, in 2011 Spring-8 Angstrom Compact free-electron Laser (SACLA) was finished and started experiments in 2012. SACLA provides a slower pulse rate (60 Hz), and ten time lower flux than LCLS (Ishikawa et al., 2012), but otherwise has many of the same parameters allowing for data collection at 60 Hz repetition rate and at a broad

wavelength range from 1-70 Å (Ishikawa et al., 2012). The PAL XFEL (Han et al., 2012) in South Korea, which will start lasing in 2016, will most likely be the third operational high energy X-ray FEL in the world. It will be followed by the Swiss X-ray FEL (SwissFEL) (Institute, 2014), which is under construction at the Paul Scherrer Institute, and will provide a very similar source to both LCLS and SACLA, with a repetition rate of 100 Hz, and a wavelength range from 1 -70 Å. The European X-ray FEL (EXFEL) will be the most advanced of the high energy FELs when it will start accepting users in 2017 (Altarelli and Mancuso, 2014). It is planned to have significantly higher flux (three times more than LCLS) than current facilities. While the energies will be slightly higher, and the wavelength range will be similar, and the major improvement the EXFEL brings is the high repetition rate of 27,000 Hz, which is a huge difference compared to current repetition rates at other FELs. These are the FELs that are currently under construction, however many more are already planned. LCLS II at the SLAC campus will start construction in 2017 and should be operational in 2020-2022, and there is a Chinese FEL in the planning stage in Shanghai. With the addition of these FEL sources the amount of beamtime around the world will greatly increase, and thereby the ability to perform new and exciting experiments.

Serial Femtosecond Crystallography: a New Concept for X-ray Crystallography

Motivation for Serial Femtosecond Nanocrystallography

With macrocrystallography already so successful, having already solved more than hundred thousand structures in the 75 years that protein diffraction has been observed, what was the impetus behind the development of structure determination with

Free Electron Lasers and the development of the new method of SFX? The answer to this question is manyfold, and needs to begin with the example of the first model protein used for SFX studies: Photosystem I. It took thirteen years from the first sighting of small crystals in 1988 (Witt et al., 1988), to the ability to determine a high resolution structure based on large single crystals under cryogenic conditions (Jordan et al., 2001).
This led to the idea to develop a new method that overcomes X-ray damage in protein crystallography and would allow structures to be determined directly from the small nano/microcrystals. If such a method could be developed and structures could be achieved at atomic resolution, it could speed up the structure determination process for difficult to crystallize proteins by years or decades because crystal size optimization and determination of cryogenic freezing conditions would become obsolete. Small crystals are also more commonly observed than large crystals during simple crystal screening, so in many cases atomic resolution structure may be solved without need for lengthy optimization in order to solve a structure.

The second advantage of structure determination with FEls is overcoming radiation damage. If one could truly outrun damage then one would be able to look for the first time at many complexes without radiation damage, and even more importantly many complexes that have site specific damage could be solved without radiation damage affecting specific sites, like the metal Mn cluster in PSII.

Finally, the sample delivery method for SFX is based on a liquid jet, where nanocrystals are delivered to the FEL beam at room temperature in their mother liquor. This is a huge advantage because it allows for data collection at conditions that may

match biological conditions more closely and may allow for time-resolved crystallographic studies at room temperature in the future.

With all of these possible improvements to structural biology, however, several problems had to be overcome. The first challenge was to find new methods to grow a sufficient quantity of extremely small crystals. These new methods, which were developed in the time course of this dissertation, will be discussed in great detail later in the following chapters. The second challenge was the delivery of the crystals to the interaction with the X-ray beam at a rate that allows for collection of data at the high X-ray pulse repetition rate of 120 Hz. The last challenge was the question of how to analyze this entirely new form of data, where hundreds of terrabytes of data are generated, crystal hits have to be identified, hundreds of thousands of individual diffraction snapshots must be indexed, and accurate structure factors must be determined.

The LCLS has several unique properties that make data collection very distinct from "normal" X-ray data collection at Synchrotron Sources. First, the LCLS has an extremely high X-ray photon flux, roughly 10^9 higher than third generation synchrotron sources. This means that when focused it can literally destroy any material that is put into the beam path. Accordingly, data collection from a single mounted crystal is completely unrealistic, except if the beam is attenuated by orders of magnitude. Thus, injectors have been developed which can deliver a fully hydrated jet of crystals to the interaction point with the X-ray laser path(DePonte et al., 2008). To further enhance the method of SFX, improvements to the jet have been made to make the jet more stable, reduce flow rate to decrease sample consumption, and to make thinner jets to reduce the background noise(Weierstall et al., 2008; Weierstall et al., 2012). This new type of sample delivery

allows for the flow of a continuous stream of hydrated crystals into the interaction region introducing fresh sample to every pulse of the X-ray laser. One major problem is that this method of sample delivery consumes a large amount of sample. The jet consumes 10-20 µL/min which leads to a sample consumption of 14 to 28 mL of sample/day, with protein concentrations often exceeding 10-20 mg/ml. This massive sample consumption is difficult to sustain for many proteins that are expressed only in small amounts. Furthermore, the common method of crystallization (vapor diffusion) cannot be used to crystallize protein in these volumes, which called for the urgent need for the development of new methods to support the reproducible growth of nano and microcrystals in large volumes.

Another major problem that is caused by the sample delivery in the jet is the fact that the data sets are not collected from one large single crystal but from individual diffraction patterns of hundreds of thousands of crystals. In a synchrotron the crystal is rotated in known degrees and therefore provides the angular integration across the entire Bragg condition. In contrast, when the crystals are delivered in a jet, the crystals appear in a random orientation and this requires indexing of each pattern individually. As each snapshot represents a still image, all reflections are partials and the structure factors must be determined over thousands of diffraction patterns. Thereby, SFX depends on merging and indexing millions of patterns, terabytes of data, and each of the diffraction patterns is an ultra-thin slice through the sphere of reflections in an unknown orientation. To handle these problems a ‘Monte-Carlo’ approach has been developed (Kirian et al., 2010). Data analysis is not the focus of this work, however I would like to direct readers who are interested in more details about the data evaluation methods for SFX and especially for

the data evaluation program CrystFEL to the following references (Kirian et al., 2011; White et al., 2012). New methods of phasing have also been attempted by using the unique properties of an X-ray FEL(Spence et al., 2011).

These short paragraphs just represent a brief summary of the various aspects of X-ray FEL experiments. For a more in-depth review please see the following references(Fromme and Spence, 2011; Spence et al., 2012).

Time-Resolved Crystallography with Synchrotron Sources

Principle of Time-Resolved Laue Crystallography

Time-resolved structural biology has always been the desired conclusion of structural studies, aiming to understand the intermediate/short-lived states of a protein reaction. In the realm of crystallography, the main technique to turn to in this pursuit was Laue crystallography (Neutze and Moffat, 2012; Spence et al., 2012). Using this approach a reaction is started in a crystal at room temperature and data is then collected using a broad-bandwidth X-ray beam. There are two advantages to using a broad-bandwidth X-ray beam: the first is that by using the full spectrum the full flux produced is also utilized, and the second is that the use of the full spectrum allows the collection of full reflections rather than the partial reflections that are usually collected, all without having to rotate the crystal. Thereby, a larger part of the reciprocal space can be covered with fewer diffraction images than in standard crystallography. This ability to collect data quickly without the need for rotation of the crystal is critical when trying to do time-resolved data collection at extremely short time scales (~100 ps) (Schmidt et al., 2010).

Accomplishments of Time-Resolved Laue Crystallography

Laue Crystallography has solved multiple structures to very high spatial and temporal resolutions. Two examples of this are myoglobin, which showed dissociation of carbon monoxide (Bourgeois et al., 2006), and studies on the Photoactive Yellow Protein (PYP), where changes of light-induced protein and cofactor conformations were detected (Rajagopal et al., 2005). These results featured difference maps showing conformational changes both large and small in multiple areas of the proteins. Especially interesting is that a time resolution to about 100 ps has already been achieved and allowed for the imaging of the conformational switch of the chromophore in PYP. Interestingly, in another recent success of Laue crystallography a team of scientists reported the ability to map not only the temporal-spatial dimension, but the pH dependent dimension as well. This work provides deeper understanding of how dynamics progress not only with time, but also as a function of the pH in regard to the possible conformational states (Tripathi et al., 2012).

Challenges of Time-Resolved Laue Crystallography

Despite the success of Laue Crystallography, there are still challenges which can make it very difficult to perform successful time-resolved Laue experiments. The first problem is that Laue has incredibly high demands on crystal quality and size (Baxter et al., 2004; Neutze and Moffat, 2012). Furthermore, the crystals must show a very low mosaicity, or internal disorder, and excellent diffraction data must be collected on the dark and excited state on one large single crystal in order to observe the changes in structure factor amplitudes. The second challenge is that frequently the changes that the researchers want to detect cause strain within the crystal, which increases the mosaic

spread (Kaminski et al., 2010). Finally, there are also some technical problems involved: there are a limited number of methods for triggering the reaction in question, and related to that is that many reactions are irreversible, which makes it impossible to repeatedly pulse and collect diffraction patterns of a single time point.

Time-Resolved SFX: Towards Molecular Movies

Motivation for Time-Resolved SFX

With the advent of the FELs and the first high energy FEL, the LCLS, which could image biologically relevant molecules, the idea to use this new ultrashort X-ray source for time-resolved studies was quickly explored. There are many reasons that an FEL provides the optimal set of tools for time resolved crystallographic studies. The first advantage is the short timing of the X-ray pulses, in the femtosecond time range, which allows for fast biological processes to be investigated that never before could be probed. This opened up many new processes for structural and functional investigations that could never have been imaged previously. Because of the injection method, irreversible processes that before would have been very hard to image at a synchrotron are also opened up as possible targets that can now be visualized. Possibly the most important motivation behind using TR-SFX to study molecules would be the damage-free diffraction of SFX data collection. Most of the interesting chemistry in proteins takes place at metal clusters, or other easily damaged cofactors embedded into the protein. With Laue, the damage to these areas (for example a metal-center) quickly accumulates so that it is impossible to determine a native structure of the protein (Kaminski et al., 2010; Spence et al., 2012). However, with TR-SFX it would be possible to get a native

structure from a stream of nanocrystals and still visualize the conformational changes that take place in the protein. At the moment of conception, TR-SFX was still limited to reactions which can primarily be photo-activated, i.e. photosynthesis, or induction by caged substrates, however work almost immediately started on injectors that would be capable of mixing the enzyme and substrate together on a relevant time-scale just before imaging. While at the moment of this writing these injectors still are not available, in the future they will open up the possibility to perform TR-SFX on many new and highly interesting proteins which before would never have been open to time-resolved structural studies.

The first proof-of-principle TR-SFX study was performed at the beamline CXI at LCLS. In the experiment, a light driven conformational change was induced in Photosystem I-Ferredoxin co-crystals (PSI-Fd) using an optical laser. PSI-Fd is an electron transport complex, where PSI gathers light and transports electrons to Fd. Upon electron transfer from PSI to Fd, the reduced Fd dissociates away from PSI to bind and deliver the electron to the Ferredoxin:NADPH reductoase. The electron transfer reaction between PSI and FD is multiphasic, with two faster time constants of 500 ns and 2 µs (Aquila et al., 2012). This complex was chosen for TR-SFX, because of the irreversible nature of the conformational change. When we shine light on the crystals, a rapid conformational change occurs, before the crystal is subsequently dissolved, making it impossible to use other traditional forms of time-resolved crystallography.

The main difference between the TR-SFX PSI-Fd experimental setup and the normal SFX experiments was the addition of the pulsed optical laser that hits the liquid jet to excite the crystals. A scheme of this experiment is shown in Fig. 1.5. Diffraction

patterns were collected at two different time delays, 5 µs and 10 µs after the optical laser pulse. Furthermore, data of the dark state were collected. Due to limitations in beamtime, unfortunately insufficient data has been collected to solve the structure, however changes in structure factors were observed and conclusions could be drawn, as seen in Fig. 1.6. A noticeable increase in structure factor amplitudes was detected between the dark state and the 5 µs time point. The other piece of information that was immediately noticed was that the 10 µs time-delay had noticeably less signal and well diffracting crystals than the other

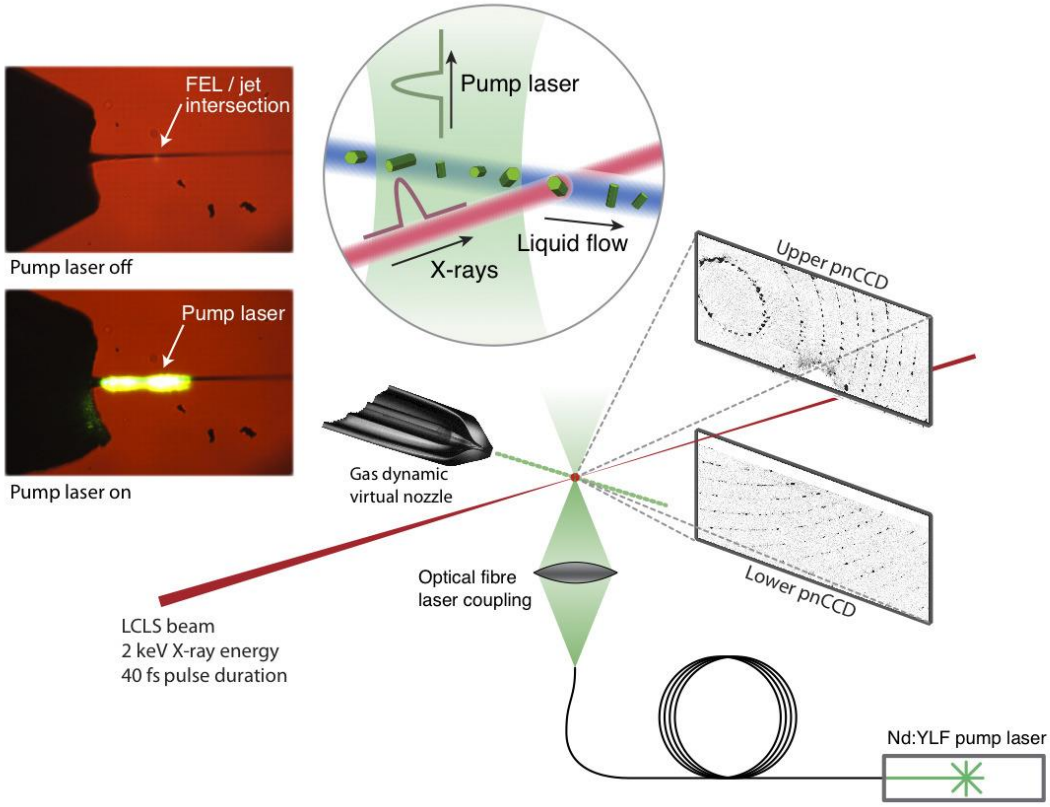


Figure 1.5. ***Experimental Set-up of the PSI-Fd Conformational Change Study****. The upper left boxes show the thermal glow of the X-rays interacting with the liquid jet. This shows the overlap of the X-ray pulse with the pump optical laser. The pump laser excites upstream in order to compensate for the roughly 130 µm travel of the crystal between the pump and probe pulses due to time-delay. Figure and caption taken with changes (Aquila et al., 2012).*

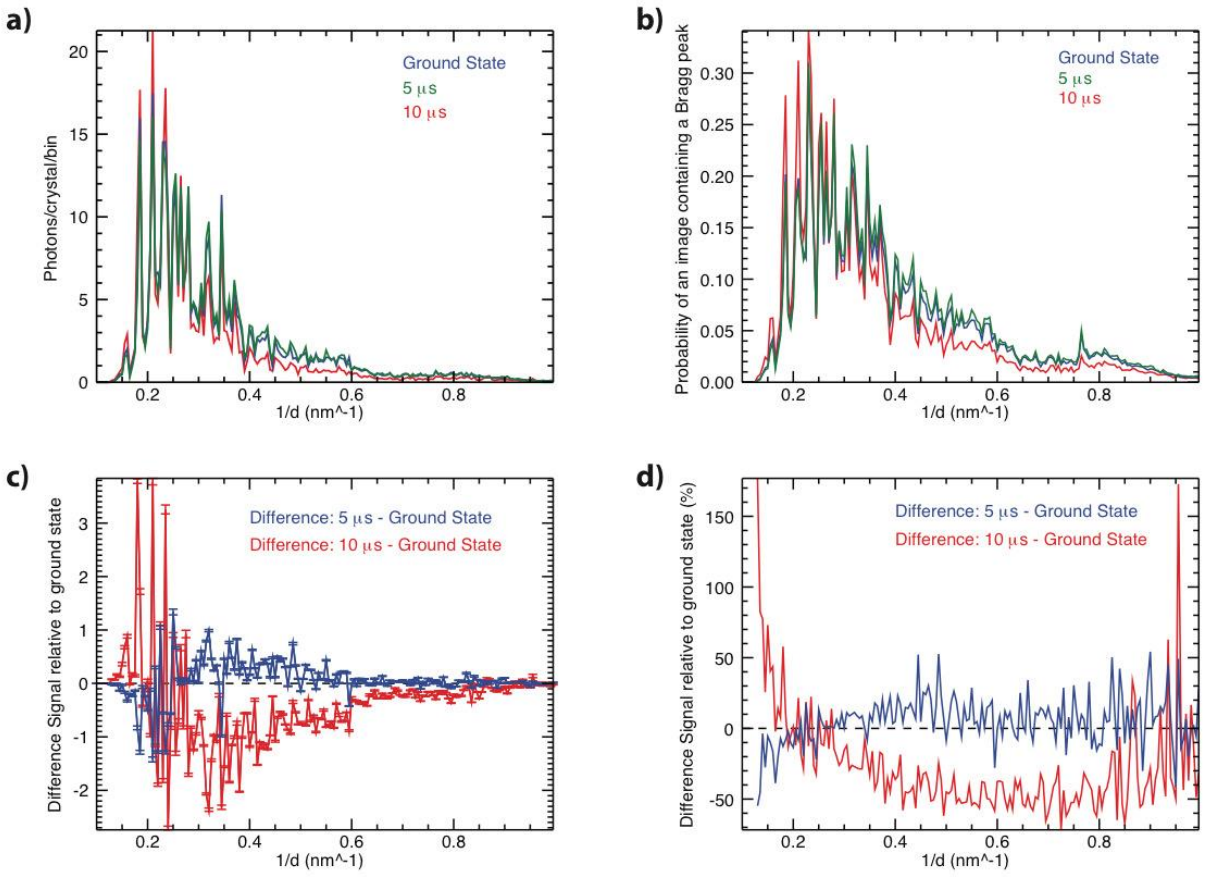


Figure 1.6. ***Data Analysis of PSI-Fd Conformational Change Study.*** *(a) The 1D virtual powder patterns for the ground state and two positive time delays for PSI-fd co-crystals. The intensities were linearly scaled to minimize the signal difference between for the different excited delays to the ground state. The powder patterns show the average number of photons scattered per crystal hit with in a histogram bin size of q = 0.005 nm−1. The area under the curve corresponds to 393.3, 406.6, and 378.1 average diffracted photons per crystal detected for the dark, 5 μs, and 10 μs data sets respectively. (b) Counting statistics of the number of peaks used in the virtual powder pattern recorded as a function of 1/d. The area under the curve corresponds to the average number of peaks per crystal diffraction pattern of 10.8, 11.5, and 9.9 peaks/pattern for the dark, 5 μs, and 10 μs data sets respectively. (c) Relative difference signal between the ground state and the excited states as a function of resolution. Error bars on the data points are 3σ errors. (d) Difference signal expressed as a percentage change. Figure and caption taken from (Aquila et al., 2012).*

two time points. These two effects culminate in a reaction cascade that leads to the dissolving of the crystals 10 µs after light exposure. These findings fit in well to the spectroscopy data that had been published on the PSI-Fd complex. (Aquila et al., 2012)

Photosynthesis: How Can SFX Help Us to Understand the Molecular Mechanism of Light Conversion and Electron Transfer?

Overview: Principles of Photosynthesis

Photosynthesis is performed by plants, algae, and bacteria, and is the most important biological process on earth. It converts light energy into chemical energy, and forms the basis for all higher life on earth. Photosynthesis can be broadly split into two categories, oxygenic (performed by plants, algae and cyanobacteria) and anoxygenic (performed by green sulfur bacteria, green non-sulfur bacteria and heliobacteria). While the principles of light conversion are similar, oxygenic photosynthesis evolves all the oxygen in the atmosphere by using water as an electron donor. This work will focus solely on the oxygenic version of photosynthesis. The primary process in oxygenic photosynthesis of light conversion and electron transfer is catalyzed by large pigment-protein complexes embedded into the photosynthetic membrane. In these reactions many pigment-protein complexes work together with multiple cofactors. The major protein complexes which are involved in oxygenic photosynthesis and their orientation within the membrane are shown in Fig. 1.7A. This figure is a model derived from electron-density maps determined by X-ray crystallography. A lot of time and effort have been devoted by many researchers in order to solve these structures (Fromme and Grotjohann, 2008).

Light Reactions: the Photosynthetic Electron Transport Chain

In oxygenic photosynthesis the electron transfer reaction initiates at Photosystem II (PSII), which captures light with its internal chlorophyll system. The energy from this captured light is used to induce a charge separation at the center of the protein complex by the primary donor P680. After charge separation occurs, P680+ is re-reduced by electrons extracted from the metal cluster in the heart of the oxygen evolving complex (OEC) in PSII. The water splitting process requires four subsequent charge separation events, where the OEC cycles through five steps, S_0 to S_4, or the Kok cycle (Shinkarev, 2003). In each charge separation event one electron is extracted from the metal cluster of the OEC in PSII. Following each charge separation there is an oxidation event involving two water molecules to finally form O_2, four protons, and four electrons, at the oxygen evolving complex (OEC). The electrons extracted from two water molecules eventually arrive at the terminal acceptor, plastoquinone (PQ), which is a mobile charge carrier. After being reduced twice it will bind two protons, forming plastoquinol (PQH_2), which is then released into the thylakoid membrane. This process is continually repeated with PQH_2 being replaced with PQ in the thylakoid membrane (Renger, 2012) (Loll et al., 2005).

PQH_2 diffuses within the membrane and will dock to the cytochrome b6f complex. This docking releases two protons into the interior of the thylakoid (lumen) and transfers the two electrons to the b6f complex. The electrons will then reduce two molecules of either plastocyanin (PC) or cytochrome c6. An additional proton will be transferred to the lumen for two electrons, thus furthering the electrochemical gradient across the thylakoid membrane(Kurisu et al., 2003).

Both PC and cytochrome c6 are soluble electron carriers which transfer the electron to Photosystem I (PSI). PSI uses light energy for the second charge separation event that transfers electrons from the lumenal to the stromal (outside the thylakoid membrane) side of the protein where the electron is used to reduce ferredoxin. Ferredoxin is a single electron carrier protein that then undocks from PSI and eventually binds to the ferredoxin: NADP+-reductase (FNR)(Setif and Bottin, 1995). For a schematic representation of this cascade of events, see Fig. 1.7B.

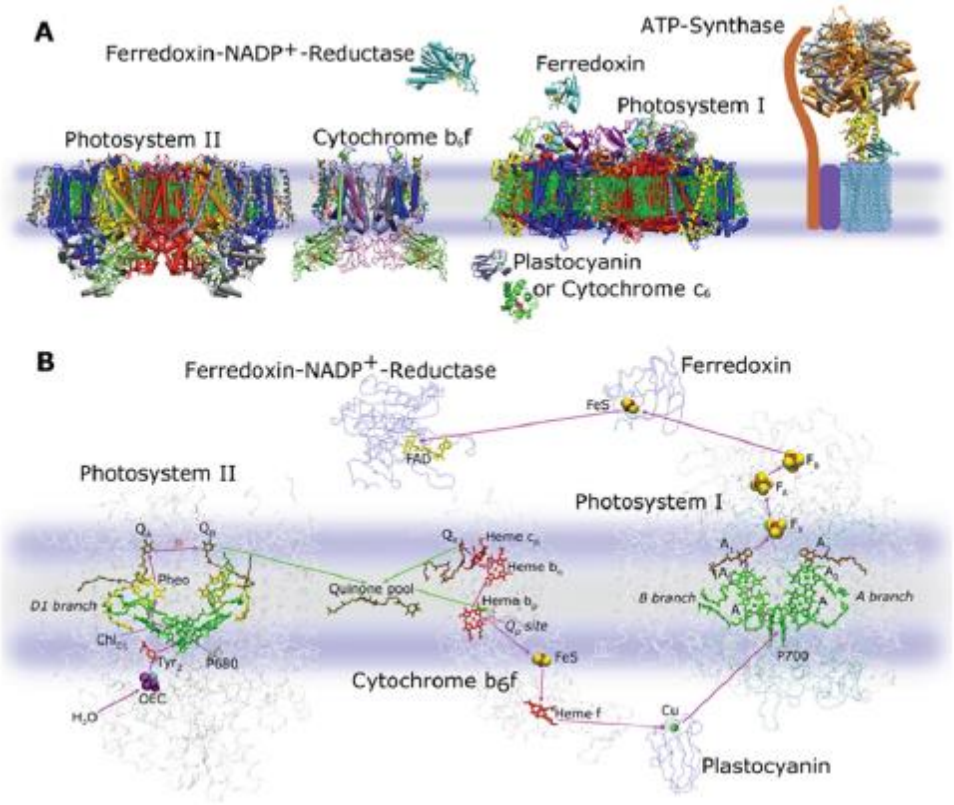


Figure 1.7. ***Overview of Oxygenic Photosynthesis****. This image shows the major protein complexes involved in oxygenic photosynthesis. (A) emphasis on the cofactors that take part in the electron transfer chain from PSII through ferredoxin-NADP+-reductase. (B) These proteins are contained within the thylakoid membranes of photosynthetic organisms, the lumen is below the membrane and the stroma is above ht membrane in these images. Image courtesy Dr. Petra Fromme.*

The electrochemical gradient that is generated by this entire process is used to produce ATP from ADP and inorganic phosphate by ATP-synthase, another protein embedded in the thylakoid membrane. Of the proteins that have been discussed here, this work will focus mainly on PSII and its oxygen evolving potential.

Structure and Function of Photosystem II

Photosystem II is the only known biological system capable of producing molecular oxygen from water and sunlight. This process provides all of the oxygen on earth, thus maintaining the biosphere. PSII is directly responsible for life on earth by maintaining the O_2 supply necessary for every other creature that relies on respiration. PSII is a large membrane protein complex. It consists of twenty protein subunits, at least fifty cofactors and is present in the native membrane, in solution, and crystalline form as a dimer. The structure of PSII was solved to 3.8 Å in 2001(Zouni et al., 2001), 3.0 Å in 2005(Loll et al., 2005) with a more complete cofactor assignment, and finally solved to 1.9 Å in 2011(Umena et al., 2011). An image based on the 1.9 Å structure (pdb code: 3ARC) is shown in Fig. 1.8. Each monomer of PSII contains four larger membrane-intrinsic subunits that are part of the electron transport chain, and also contain most of the antenna system. They are labelled: D1 (PsbA), D2 (PsbD), CP43 (PsbC), and CP47 (PsbB). The D1 and CP43 subunits are particularly relevant to the oxygen evolving complex (OEC) because they provide the ligands for coordination of the manganese cluster. Thirteen other low molecular mass, membrane-intrinsic proteins are also present, surrounding the core subunits, plus three lumenal protein subunits that stabilize the metal cluster in the heart of the OEC.

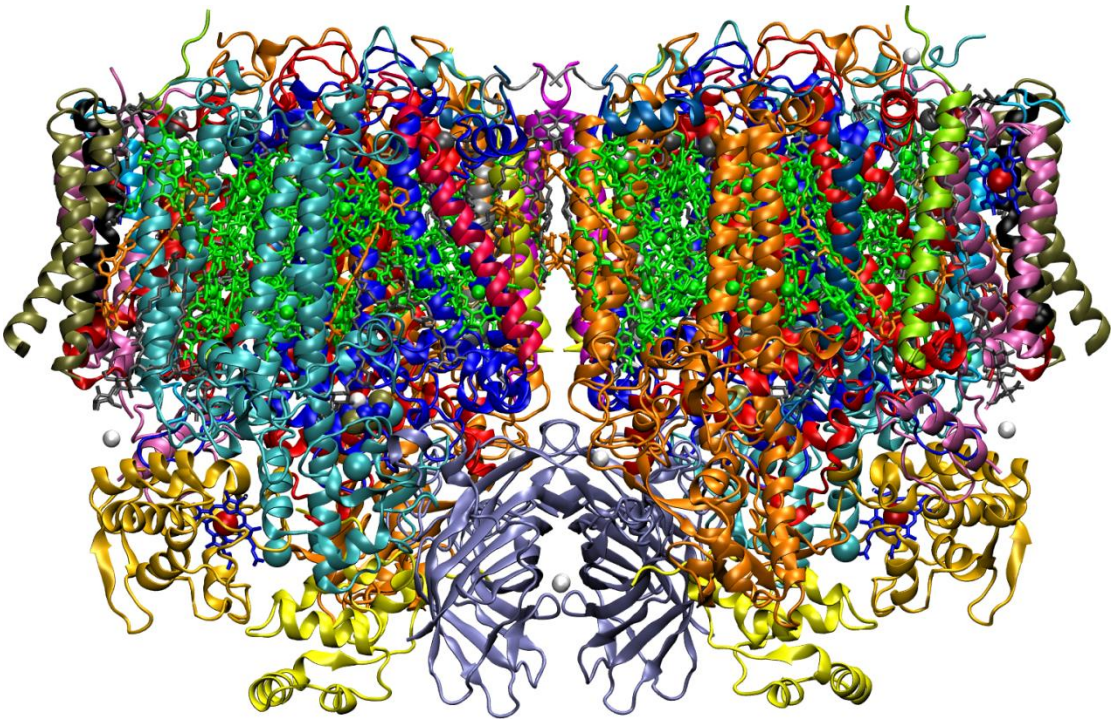


Figure 1.8. ***Overall Structure of PSII Dimer.*** *The protein subunits are colored individually in the right hand monomer Image based on the structure of pdb code: 3ARC (Umena et al., 2011).*

One PSII monomer contains 36 transmembrane (TM) helices. Ten of these helices are part of the D1 and D2 subunits. The CP43 and CP47 subunits account for another twelve TM helices. The membrane loop regions of D1, D2, CP43, and CP47, along with PsbO, PsbU, and PsbV, which are the three luminal subunits, house the OEC (Umena et al., 2011).

The electron transport chain of PSII is described briefly below, but, in short, is comprised of four chlorophyll a, two pheophytins, and two plastoquinones, one immobile and one mobile. Interestingly enough, only one branch of this chain is actually active, which is the D1 branch, see Fig. 1.9. The charge separation occurs at the quartet of

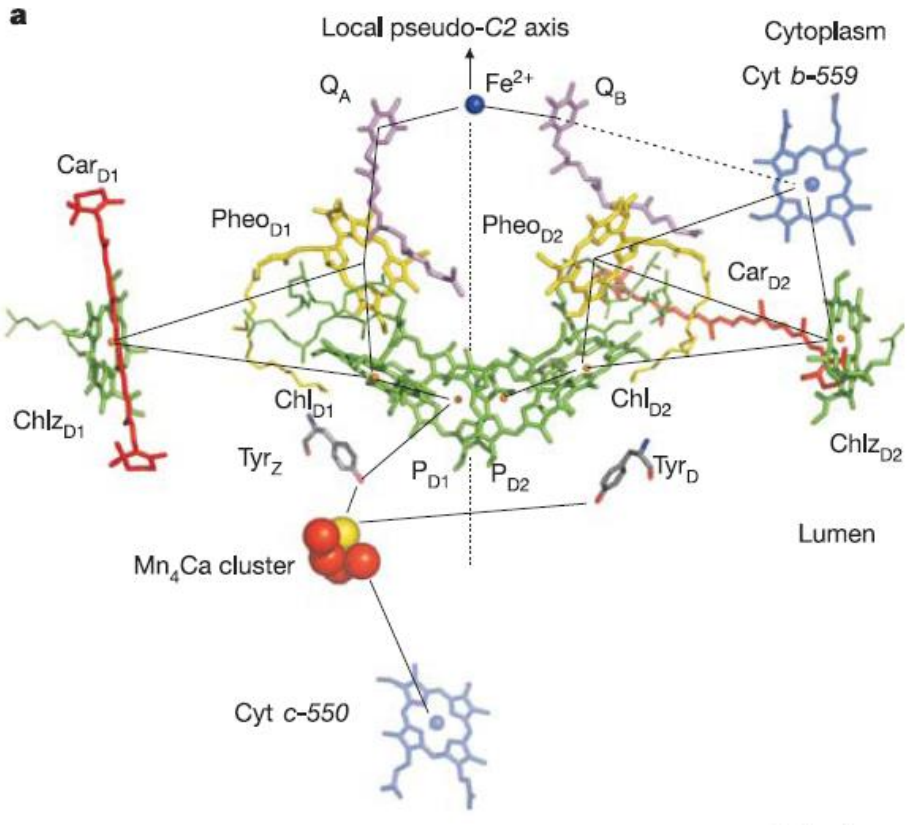


Figure 1.9. ***Electron Transport Chain of PSII****: This view is parallel to the membrane plane. The electron transport chain is arranged into two branches, the left branch (D1 branch) and the right branch (D2 branch). Only the D1 branchis active. Image and caption taken with changes from (Loll et al., 2005).*

chlorophyll molecules called the P680. After that the electron follows the D1 side to the pheophytin, then the PQA, and finally arrives on the PQB binding site. Once two charges have gathered, forming PQB^{--}, it binds two protons, forming PQH_2, which is then released into the membrane. The $P680^+$ that is initially formed is reduced by a tyrosine that donates an electron to $P680^+$. This tyrosine is located between the P680 and the OEC (Zouni et al., 2001). The OEC catalyzes the water oxidation process of PSII, and is the most interesting part of the whole complex. The OEC is comprised of a Mn_4CaCl metal cluster. This center couples water oxidation, a four electron extraction process, to the one

electron photochemistry used by the reaction center. This occurs due to the ability of the manganese atoms to store charges as oxidation states following the Kok cycle (Renger, 2012), S_0 to S_4, as previously described. Each charge separation event by P680 extracts one electron from the OEC, which after four separation events leads to the evolution of one oxygen molecule, thus resetting the whole system to start again.

Structure and Function of Photosystem I

PSI is a large membrane protein complex that uses light energy to catalyze a charge separation event. In this light-driven reaction electrons are transferred from the luminal to the stromal side of the thylakoid membrane. These electrons are then eventually used to reduce $NADP^+$ to NADPH. PSI has multiple oligomeric forms, in plants it exists as a monomer bound to four light harvesting proteins (LHCI), while in cyanobacteria it is usually a trimer (Fromme and Grotjohann, 2008), but when grown in high light there can be a mixture of both trimer and monomer. PSI from cyanobacteria is the largest and most complex membrane protein that has been crystallized(Fromme and Witt, 1998) and from which a crystal structure at high resolution has been solved (Jordan et al., 2001).

Each monomer of cyanobacterial PSI consists of twelve protein subunits and 127 non-covalently bound cofactors. The cofactor composition consists of 96 chlorophylls, 22 carotenoids, three Fe_4-S_4 clusters, three lipids, two phylloquinones, and one Ca^{2+} ion (Jordan et al., 2001). A model of trimeric PSI is shown in Fig. 1.10. The main subunits of PSI are PsaA and PsaB, which form a heterodimer at the center of the monomer. This heterodimer is both a reaction center and a core antenna complex. Both of these subunits contain eleven TM helices and together they coordinate with over half of the chlorophyll

molecules and most of the carotenoids (Jordan et al., 2001). In addition to this, the heterodimer coordinate most of the electron transport cofactors including P700, A, A0, A1, and the iron-sulfur cluster Fx.

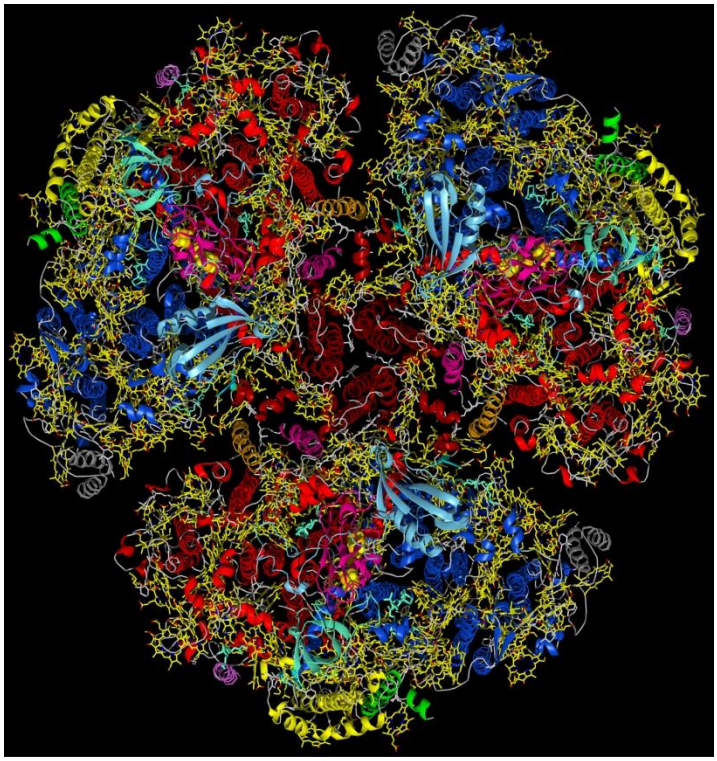


Figure 1.10. ***Model Structure of PSI Trimer.*** *This is a view normal to the membrane of trimeric PSI (pdb code: 1JBO). This is model is to a resolution of 2.5 Å (Jordan et al., 2001). Image courtesy of Dr. Petra Fromme.*

The core of PSI coordinates seven protein subunits inside of the membrane: PsaF, PsaI, PsaJ, PsaK, PsaL, PsaM, and PsaX. Three of these are involved in the trimer contact sites: PsaI, PsaL, PsaM, while the others are in the membrane-exposed region of the monomers. The stromal hump, the area of PSI that extends out into the stroma, is made up of three protein subunits: PsaC, PsaD, and PsaE. PsaC holds F_A and F_B, the terminal

two iron-sulfur clusters, PsaD is necessary for assembling, and PsaE may also be involved in the binding site of ferredoxin (Fischer et al., 1998).

The PSI electron transport chain is comprised of six chlorophyll molecules, two phylloquinones, and three iron-sulfur clusters (Fe_4-S_4). Traditionally it was believed that charge separation is induced using light energy at the P700 special pair located at the center of a PSI monomer (Fromme and Grotjohann, 2008), however recent experiments have shown that instead charge separation may be initiated at one of the two accessory chlorophyll molecules (Muller et al., 2010). This electron is transferred via A, a chlorophyll a molecule, to A0, another chlorophll a, and then to A1, a phylloquinone, see Fig. 1.11 for illustration. From here the electron is shuttled to the iron-sulfur clusters via the pathway F_X, F_A and the terminal cluster F_B. From F_B, ferredoxin takes the electron up to bring it to FNR, which catalyzes the reduction of $NADP^+$ H^+ to NSADPH (Fromme et al., 2003; Ceccarelli et al., 2004).

While many of these electron transfer pathways are understood fairly well, further structural and functional studies could be of extreme value for the understanding of the molecular mechanism, which would explain exactly how this amazing set of protein complexes manage to keep everything on earth alive.

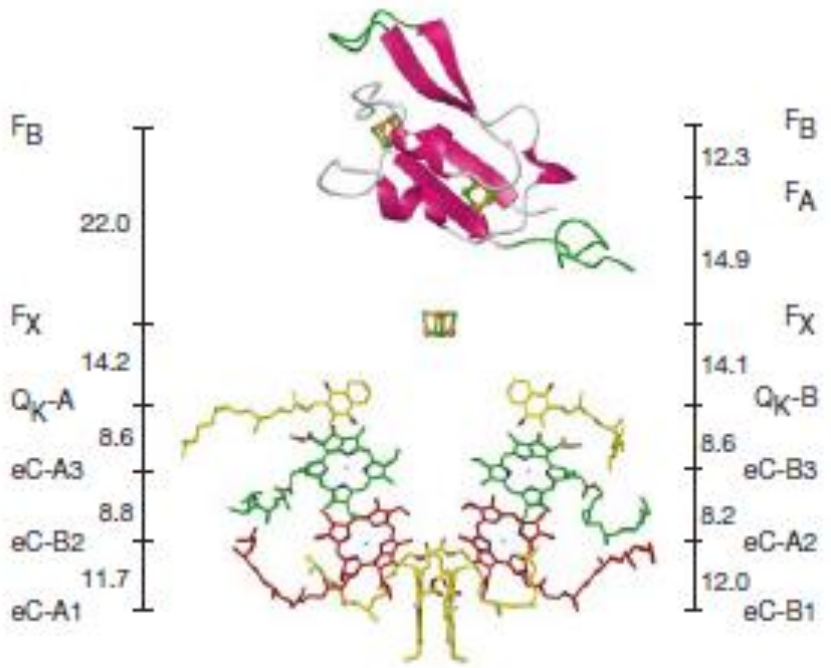


Figure 1.11. ***Electron Transport Chain of PSI****: This view is parallel to the membrane plane. The electron transport chain is arranged into two branches, the left branch (A branch) and the right branch (B branch), denoting either PsaA or PsaB. The label eC stands for chlorophyll and Q_k represents the phyloquinone. Three Fe_4S_4 clusters are located along the stromal side of the protein. Image and caption taken with changes from (Jordan et al., 2001).*

How SFX Can Help us to Understand the Molecular Mechanisms of Photosynthesis

SFX could truly push our understanding of photosynthesis to a new level. Through the use of SFX and TR-SFX, a more thorough understanding of the process of oxygen evolution and electron transfer could be gained, by providing the first ever truly native dark state of PSII, as well as making it possible to visualize the PSI-Fd interaction and movement upon reduction. Thereby SFX would make a perfect model system for electron transfer between separate proteins. Also, large complexes like PSI and PSII naturally contain many subunits, cofactors, and metal atoms inside of them. By being able to take an image of these without radiation damage we can gain a more complete

understanding of the process of photosynthesis, which is arguably the most important biological process supporting life on this planet.

Motivation of this Dissertation

Development of New Methods for Nanocrystal Growth and Characterization

Since the discovery of the basic principles of crystallography at the beginning of the 20th century, the goal of X-ray crystallography has been to grow the largest and most organized crystals that are possible, with the aim to solve the structure of the protein. The advent of SFX has caused a complete paradigm shift from trying to grow the largest crystals possible to attempting to grow the smallest most reproducible crystals possible, while obtaining a high enough yield from the target protein to sustain the extremely high sample consumption rate.

This was a major challenge of the thesis at the start, because none of the traditional techniques for crystallography were still applicable. Even if a shower of nanocrystals was obtained from a screen, it was unrealistic to attempt to collect 3 μL drops in order to supply 7.2 mL of protein crystal suspension per day. Therefore, new techniques had to be created in order to allow SFX to be successful as a protein structure determination method.

Also with the need of nanocrystals for SFX another necessity was born – the ability to determine the difference between crystalline and non-crystalline material on a size limit that is at the diffraction limit of many optical microscopes. To this end Second Order Non-linear Imaging of Chiral Crystals (SONICC) and dynamic light scattering (DLS) have been developed as characterization tools for SFX.

Understanding Photosystem II with the Development of TR-SFX

PSII is one of the most complex membrane proteins ever solved. This is secondary however to the fact that it is arguably a significant part of the most important biological process on Earth, the splitting of water in hydrogen and oxygen that allows all life on earth to survive. Understanding how this process works has been the goal of many scientists' careers. Unfortunately, because of its extreme sensitivity it is not truly capable of being solved by more traditional time-resolved methods. This is mainly due to the susceptibility of the manganese cluster to radiation damage. Once the cluster is damaged it is in an unnatural state that it cannot recover from and is therefore significantly less helpful for determining the conformations during which water is split.

TR-SFX, however, is the perfect technique for solving this problem in a new and unique way. There is no radiation damage, the crystal supply is constantly refreshed, and because of the small size of the crystal it is possible to excite the entire crystal at once, not just a small portion of the crystal. This makes TR-SFX an extremely exciting technique for attempting to solve the quandary that is PSII and oxygen evolution.

SFX Techniques for Multiple Protein Crystallization and Challenging Membrane Proteins

Membrane proteins have long been a very difficult subset of proteins for solving structures. There are many reasons behind this, but one of the primary ones is that they contain both hydrophobic and hydrophilic portions. Due to this it is very difficult to grow a crystal large enough to solve the structure using macromolecular crystallography. SFX

gives the ability to use these small crystals for structure analysis, however requires a large volume of crystal suspension. The new techniques described here may initiate a new field of X-ray crystallography.

Chapter 2

Microcrystallization Techniques for Serial Femtosecond Crystallography Using Photosystem II from *Thermosynechococcus elongates* as a Model System

Kupitz, C., Grotjohann, I., Conrad, C.E., Roy-Chowdhury, S., Fromme, R., Fromme, P.

Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604 USA

Serial femtosecond crystallography (SFX) is a new emerging method, where X-ray diffraction data are collected from a fully hydrated steam of nano or microcrystals of biomolecules in their mother liquor using high-energy, X-ray free-electron lasers. The success of SFX experiments strongly depends on the ability to grow large amounts of well-ordered nano/microcrystals of homogeneous size distribution. While methods to grow large, single crystals have been extensively explored in the past, method developments to grow nano/microcrystals in sufficient amounts for SFX experiments are still in their infancy. Here, we describe and compare three methods (batch, free interface diffusion (FID) and FID centrifugation) for growth of nano/microcrystals for time-resolved SFX experiments using the large membrane protein complex Photosystem II as a model system.

Introduction

X-ray crystallography is the most prolific technique for solving protein structures in structural biology. Structure determination of soluble proteins has made great progress in recent years, with more than 90,000 structures solved. However, the structure determination of difficult to crystallize proteins, like large multi-protein complexes and membrane proteins is severely lagging behind, with less than 400 unique membrane protein structures determined to date (Moraes et al., 2014). One of the rate-limiting steps

for the structure determination of membrane proteins with standard crystallographic methods is the growth of large well-ordered single crystals. The determination of membrane protein structures solved to date often involved a long process taking years (or sometimes even decades) to grow large, well-ordered crystals suitable for X-ray structure determination. X-ray damage is a major problem in standard X-ray crystallography for many protein crystals (Schmidt et al., 2012), especially when they contain redox active cofactors (Yano et al., 2005), therefore imposing a limitation for X-ray diffraction on microcrystals, even under cryogenic conditions (Burmeister, 2000).

The new method of serial femtosecond crystallography (SFX), overcomes many of the limitations of conventional X-ray crystallography (Chapman et al., 2011; Fromme and Spence, 2011; Spence et al., 2012). It is remarkable that the first proof of principle for SFX was done not with lysozyme or any other small, easy to crystallize protein but with Photosystem I (PSI), which consists of 36 proteins to which 381 cofactors are non-covalently bound (Chapman et al., 2011). In SFX, tens of thousands of diffraction patterns can be collected in minutes on fully hydrated nano and microcrystals in their mother liquor, at room temperature. The X-ray laser pulses are so short that they 'outrun' X-ray damage by the diffract before destroy principle (Barty et al., 2012) which also opens new avenues for time-resolved crystallography (Aquila et al., 2012; Neutze and Moffat, 2012). Unlike traditional crystallography, the SFX technique delivers thousands of small crystals of micrometer size in a liquid stream to femtosecond X-ray pulses, using an X-ray free-electron laser (XFEL). This technique is advantageous because smaller crystals have less long-range disorder compared with larger crystals and smaller crystals can be easier to grow. The decrease of long-range disorder in nanocrystals is particularly

advantageous for crystals of membrane proteins and large protein complexes, which are often the most difficult to crystallize. These crystals are often plagues by long-range disorder and anisotropic resolution leading to high mosaicity. Before SFX, the growth of small crystals was undesired, making their growth and characterization relatively unexplored. The first study of crystals grown specifically for SFX described the growth of nano and microcrystals of PSI. In this study, the crystals were grown by ultrafiltration at low ionic strength. In the ultrafiltration method, the protein solution is brought into the nucleation zone by slowly concentrating the protein at low ionic strength. This method can also be used for crystallization of proteins at high ionic strength or using other precipitants that pass through ultrafiltration membranes (unpublished data). It is however not suitable for crystallization of membrane proteins when using higher molecular weight polyethylene glycols (PEG), as PEG is a very flexible and elongated polymer that does not easily pass through ultrafiltration membranes (even PEG 1,000 with a MW of only 1 kDa does not pass quantitatively through UF membranes with a 1,000 kDa cuff-off). Unfortunately, PEG is one of the most commonly used precipitants in crystallography, and the majority of membrane proteins, including Photosystem II (PSII), (Zouni et al., 2000; Kern et al., 2005; Shen et al., 2011) have been crystallized in the presence of PEGs. This work explores PSII, a large membrane protein complex to explore techniques for nano/microcrystal growth and characterization.

PSII is one of the most important enzymes in the process of photosynthesis, converting light energy from the sun into chemical energy. It catalyzes the light-driven, transmembrane electron transfer from water, which serves as the electron source, to the plastoquinone. The reaction of water oxidation involves four charge separation events,

where two water molecules are oxidized, leading to the generation of four protons, four electrons and the evolution of one molecule of oxygen (for a review on PSII see (Renger, 2012)). This reaction is unique in nature, where PSII has produced all the oxygen in the atmosphere and has changed our planet from an anoxygenic to an oxygenic atmosphere 2.5 billion years ago. This reaction makes PSII an attractive candidate for renewable energy and therefore, understanding its structure and function are vital (Barber, 2009; Gust et al., 2009). However, PSII is an extremely difficult protein to work with. The results of our time-resolved structural work on PSII, which was presented at the Royal Society Workshop, will be published elsewhere (Kupitz et al., 2014). Here, we focus on method developments for the growth of nano/microcrystals, using PSII as a model system. The techniques presented will have a broad impact on current and future projects that aim to determine the structure and dynamics of biomolecules using SFX.

The method of SFX and the success of the experiments at XFELs depends on the growth of high-quality nano or microcrystals that are delivered to the FEL beam in a liquid stream of their mother liquor. Most SFX experiments have been performed using an injector with a virtual gas-dynamic nozzle (DePonte et al., 2008; Weierstall et al., 2012), which is installed at the CXI beamline at LCLS and delivers the sample with a flow rate of 10-20 µl min^{-1}. Alternate injector designs, with lower flow rates have also been developed. However, they require proteins to be delivered in highly viscous media. They include the injector described by Sierra et al. (Sierra et al., 2012) that delivers proteins by electro spinning and the recent development of a novel injector design that allows delivery of protein crystals in lipidic cubic phase (Weierstall et al., 2014).

For SFX on protein crystals, large quantities of nano/microcrystals of uniform size and quality are desired. While method development of crystal growth for standard X-ray crystallography has been largely focused on new methods to increase the size and quality of the crystals, methods for growth of high-quality nanocrystals are highly desired, yet largely unexplored.

We use PSII as a test case to compare different methods of nanocrystal growth and show how diagnostic tools like dynamic light scatters (DLS) and second-order nonlinear imaging of chiral crystals (SONICC) (Wampler et al., 2008) can be used to optimize the growth of nanocrystals for SFX.

Material and Methods

Cell Growth and Protein Purification

Currently, the method of SFX, using a liquid injector, requires large amounts of protein. The typical SFX experiments are performed at a flow rate of 10-20 µl min^{-1} and microcrystal suspensions with 10^{10}-10^{11} crystals ml^{-1}, depending on the desired 'hit rate.' The concentration of protein for collection of the SFX data hereby depends on the size distribution of the crystals, and for 1-5 µm crystals is commonly in the range of 10-20 mg ml^{-1} (Chapman et al., 2011; Boutet et al., 2012; Redecke et al., 2013). Thereby, the amount of protein required for SFX experiments is large and one 12 hour experiment can easily require 50-100 mg of protein.

PSI and PSII were one of the first proteins used as model systems for SFX (Chapman et al., 2011; Kern et al., 2012), and it is often assumed that they were chosen because they are abundant in nature, can be isolated in large amounts, and are easy to

isolate and crystallize. Unfortunately, these assumptions are wrong. The photosystems have been used as model systems because of their extreme importance for bioenergy conversion on the earth despite the fact that they are probably one of the most difficult proteins to work with. PSII is a large multiprotein cofactor complex, consisting of 19 proteins and more than 50 cofactors (Umena et al., 2011). It undergoes constant remodeling in the native membrane due to photodamage (Takahashi and Badger, 2011). In the native membrane, PSII has a half-life of 30 min, with between 30 and 70% of PSII being under reconstruction in the native membrane depending on the environmental conditions like light, temperature, pH, cell density, etc. it is therefore of extreme importance to grow the cells under reproducible conditions. We grow the cells of *Thermosynechococcus elongates* at 56ºC in Casteholz medium in a custom-designed 1221 photobioreactor (shown in Figure 2.1), which controls all important parameters. It increases the light intensity in parallel to the cell density and controls the pH by feeding an air/CO_2 mixture into the system. The cells are grown in a continuous mode, therefore keeping them in the logarithmic growth phase, and 301 of the cell culture are harvested each week, which corresponds to 22-28 G wet cell mass. Compared to cultures of *E. coli*, the yield of cells is low with less than 1 g of cells harvested per litre of cell culture.

Figure 2.1. ***Photobioreactor*** *developed for large-scale growth of Photosynthetic algae and cyanobacteria. The reactor has a capacity of 122 l. It can be* in situ *sterilized and allows for control of light intensity, cell density, temperature, pH and gas flow.*

The purification of PSII from the cells is performed in principle as described by Zouni et al. (Zouni et al., 2000) that includes isolation of photosynthetic membranes by differential centrifugation, solubilization of PSII with the detergent beta-dodecylmaltoside and purification of PSII by ion0exchange chromatography. In our modidfication of the Zouni et al. protocol, cell disruption is performed using a microfluidizer, which breaks the cells rapidly by shear forces in a small capillary at

18,000 psi in the flow through-mode, where 80 g of cells can be broken quantitative in less than 2 min on ice. Also, large-scale column purification is performed in a staggered mode to increase capacity. Immediately after column purification three to four re-crystallization steps (see below) are done to further purify the protein. The total yield of PSII in the form of small crystals 40-50 mg PSII per 22-28 g cells corresponding to 301 culture, which only corresponds to 1.3-16 mg protein crystals per litre cell culture, i.e. (2.5mM Chl) corresponding to 10-20 protein preparations are required to produce the sufficient nanocrystals for five shifts of TF-SFX experiments.

Crystallization as a Last Purification Step

Only fully active dimeric PSII with intact subunit composition crystallizes, while PSII in the process of assembly/disassembly is missing all or part of the luminal extrinsic proteins PsbV, PsbU or PsbO does not crystallize. This occurs because PsbV is involved in crystal contacts therefore crystallization can be used as a last purification step. In this process, PSII is precipitated three times at a Chl concentration of 1.74 mM, with decreasing concentrations of precipitant Buffer D_x (100 mM Pipes pH 7.0, 5 mM $CaCl_2$, 5 mM $MgCl_2$, X % PEG2000), 7.5%, 6.5%, and 5.5% respectively. The solubility decreases as homogeneity of the protein sample increases; therefore, the PEG concentration is decreased in each of the re-crystallization steps. The first and second precipitation/crystallization steps are allowed to progress for 1 h in complete darkness, on ice. The third precipitation is allowed to proceed for 8 h under the same conditions on ice.

Comparison of Different Methods for the Growth of Nanocrystals

Three different techniques are compared to obtain microcrystals for SFX: the batch method, free interface diffusion (FID) and FID with centrifugation. All experiments are carried out at a chlorophyll concentration of 0.5 mM Chl at 10ºC aiming for nanocrystal growth in 24-48 h.

Batch Method

To grow crystals in a batch experiment, the phase diagram should be known, so as to ensure that the solution is in the nucleation zone after mixing. The determination of the solubility curve can be quite time-consuming and is best performed by stepwise dissolving of crystals by lowering the precipitant concentration. This can be done either manually or using a crystallization robot by automated stepwise dilution of the precipitant solutions in the reservoirs. We determine the borderline between metastable and nucleation zone by small-scale batch experiments, where seeding crystals are added. In the nucleation zone, addition of seeding crytals leads to massive secondary nucleation. By contrast, the seed crystals added to the metastable zone grow without new crystals appearing. A schematic phase diagram indicating suitable starting points for batch crystallization experiments is shown in figure 2.2.

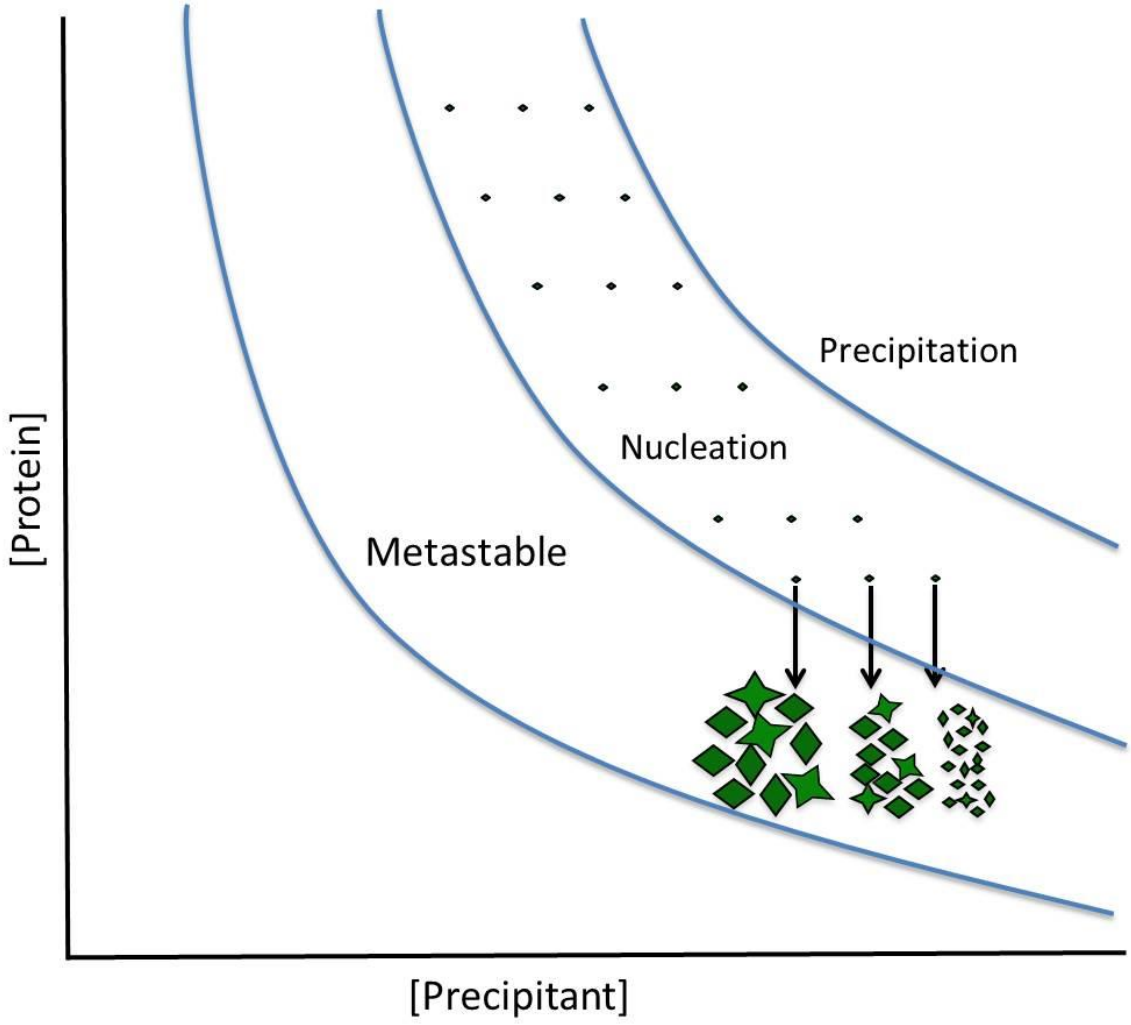


Figure 2.2. ***Schematic phase diagram*** *with suitable starting points for batch experiments. All batch experiments should ideally start in the nucleation zone. The nucleation rate increases with the increase of supersaturation, leading to a larger amount of smaller crystals.*

Once the phase diagram has been determined, the batch experiments are performed by simple rapid mixing of the protein with the precipitation buffer, under conditions where the solution will be in the nucleation zone after mixing. The batch crystallization is routinely tested at five different protein concentrations combined with a fine screening of precipitant concentration. This is performed in small scale before the

complete protein sample is crystallized in a large batch. The result of such a batch experiment with PSII is shown in figure 2.3.

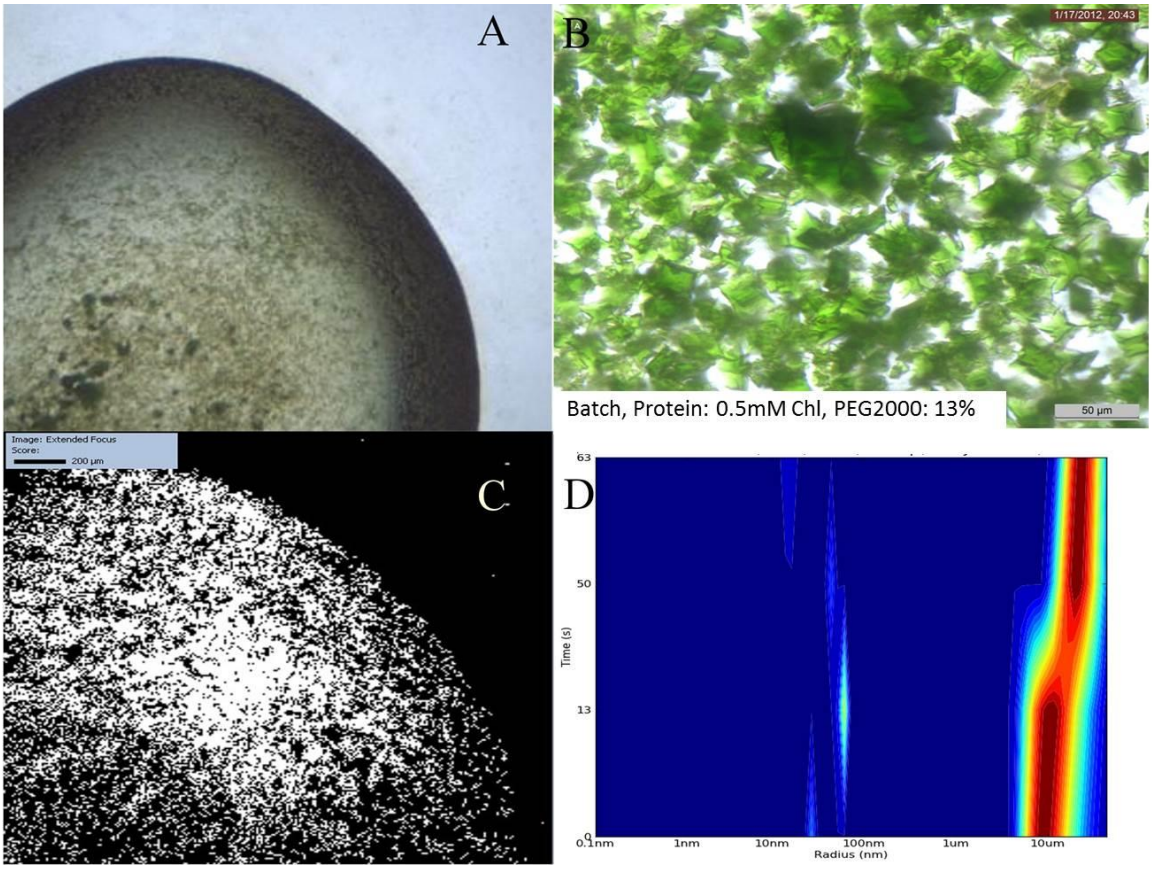


Figure 2.3. ***Batch method experiment*** *performed using 0.5 mM Chl, and 13% PEG2000 as starting conditions.(a,b) Images of crystals created using the Batch Method, there are numerous large crystals, and these are obviously poly-crystalline. (c) The SONICC image of the entire drop of (b) confirming their crystallinity. (d) DLS histogram showing that the majority of the crystals are around 10 µm in radius.*

Seeding is regularly used for growth of large single crystals but is also extremely helpful for growth of nanocrystals in batch. Addition of small nanocrystals (best less than 500 nm) to the precipitant solution prior to mixing of precipitant and protein, leads to a massive increase in nucleation rate. It also can overcome a serious problem that is observed for nanocrystal growth where proteins do not spontaneously crystallize but

crystals grow out of an amorphous precipitate in the time frame of weeks. This crystallization condition is the worst-case scenario for nanocrystal growth, but can be partially overcome by adding seeds to the precipitate. The only question could now be asked how to obtain seeds, when only larger crystals have been obtained. In this case, one can break the large crystals and use them as crude seeds in step 1 of the experiment. Once these large seeds induce nucleation, nice nano and microcrystals appear in large showers which can then be used in the next step as seeds. While all these methods help to grow nano and microcrystals for SFX in batch, the size distribution achieved is often very broad and varies between individual experiments. Furthermore, crystal growth is also very fast, so nice small crystals can grow in a few minutes into larger crystals with visible defects.

Crystal growth can progress rapidly, so crystal growth must be monitored by DLS in short time intervals to stop crystal growth at the desired size by rapid dilution in high precipitant buffer. Unfortunately, the crystals of PSII formed by the batch crystallization method tend to be poly-crystalline, resembling starfish. DLS showed that the predominant size of the crystals is 5-20 µm, but optical inspection showed large crystals up to 300 µm in size, these large crystals can be removed from the solution by pre-filtering and inline filtration. SFX data can be collected on these 5-20 µm crystals even if they are starfish-shaped. Most diffraction patterns are single crystal diffraction patterns because the small X-ray focus of only 1 µm 'hits' in most cases only one of the 'loves' of the starfish-shaped crystal. However, crystals of this size (even if they are perfectly shaped) are not suitable for TRSFX as data collection and evaluation face multiple problems: higher mosaicity of the crystals due to their larger size, non-uniform light

excitation of PSII in the crystals due to absorption of the photons within the crystal, extremely low hit rates, high sample consumption and fluctuations in the liquid jet position. The jet is only 4 µm in diameter; if the size of the crystals exceed the size of the jet, the jet becomes unstable and wiggles or even 'jumps' when a large crystal exits the nozzle. This further decreases the hit rate as the FEL X-ray beam may no 'hit' the jet when it moves. Furthermore a kinked or wiggling jet path can lead to deposition of materials on the walls of the shroud. This material accumulates at the walls of the injector forming stalagmites of the precipitants (here PEG, salts, and buffer) that slowly grow into the interaction region, whereupon data collection has to be stopped and the chamber has to be evacuated and cleaned.

Free Interface Diffusion

In this method, crystals are grown at the interface between a highly concentrated protein solution and the precipitant solution. It was designed to grow small, well-ordered crystals to overcome the problems described above for SFX data collection on 5-20 µm crystals.

FID is used in standard crystallography to create a continuous concentration gradient leading to growth of crystals of increasing sizes along a concentration gradient. Its most common use is in form of the Granada Crystallization box experiments. Recently, microfluidic devices have been developed to use FID for mapping of the phase diagram and optimization of nanocrystal growth (Abdallah et al., 2013); however, they have not yet reached the capacity for large-scale growth of 50-100 mg of nanocrystals of PSII for time resolved SFX studies. The standard methods of FID used in traditional crystallography, where large crystals are desired, use small capillaries for the

crystallization experiments to minimize the size of the interface region and create a smooth linear concentration gradient. Using this setup, one can achieve nanocrystal growth at the interface but the crystal size increases along the concentration gradient, which leads to very low yield of nanocrystals. Furthermore, it would be painstaking to accumulate 1- ml of nanocrystal suspension by harvesting nanocrystals from the interface of hundreds of small capillaries. To maximize the surface/volume, we experimented with several different setups and geometries but the best result came from the simplest method, performing the crystallization experiments in 1.5 ml reaction vessels. For the initial screening of conditions, we use volumes of 20 µl protein and 20 µl precipitant solution, which can be scaled up to 500 µl plus 500 µl. In the first sets of experiments, we tried to form a 'perfect' free interface by layering the protein solution (which has a lower density than the precipitate) very carefully on top of the precipitant (figure 2.4a). While nanocrystals formed at the interface, the yield was low and these crystals grew fast into larger starfish-like crystals similar to what is shown for the batch experiments in figure 2.3.

The nucleation rate was dramatically increased by the inversion of the set-up, shown in figure 2.4b. Here, the protein is pipetted into the bottom of the reaction vessel. This is followed by slowly dropping the precipitant solution through the protein layer at a rate of roughly 20 µl min^{-1}. This procedure causes a large transient interface area where the small drops of the precipitant move through the protein to form two layers. Eventually the two phase system forms with the precipitant solution forming a dense bottom layer with the protein on top.

The results of one of these experiments are shown in figure 2.5. The results show that the crystals are significantly smaller, with an average radius of 1-2 μm. The crystals also display much less poly-crystallinity, with most of the crystals forming a single crystal.

Free Interface Centrifugation

This third method represents a further modification of the FID method, where the formation of the two phase system is followed by centrifugation. A schematic drawing of the process is shown in figure 2.4c. Here, the reaction vessel is centrifuged after the interface was formed, at a slow speed of 200-500 rcf using a swing bucket or a fixed angle centrifuge. The centrifugation has two positive effects that support the formation of well-ordered nanocrystals: (i) the nanocrystals start sedimenting into the precipitant when they have reached a specific size. As there is no protein present in the precipitant solution, the growth of the crystals stops as soon as they enter the precipitant layer. Thereby a very uniform size distribution of small nanocrystals is achieved. (ii) Small nuclei, which form in the protein solution by diffusion of the precipitant into the protein layer, are accelerated towards the interface where the increased precipitant concentration allows them to grow into stable nanocrystals, which then enter the precipitant zone where crystal growth stops. Thereby a constant ‘stream’ of newly formed nuclei enters the interface zone and continues to grow into nanocrystals until they enter the precipitant

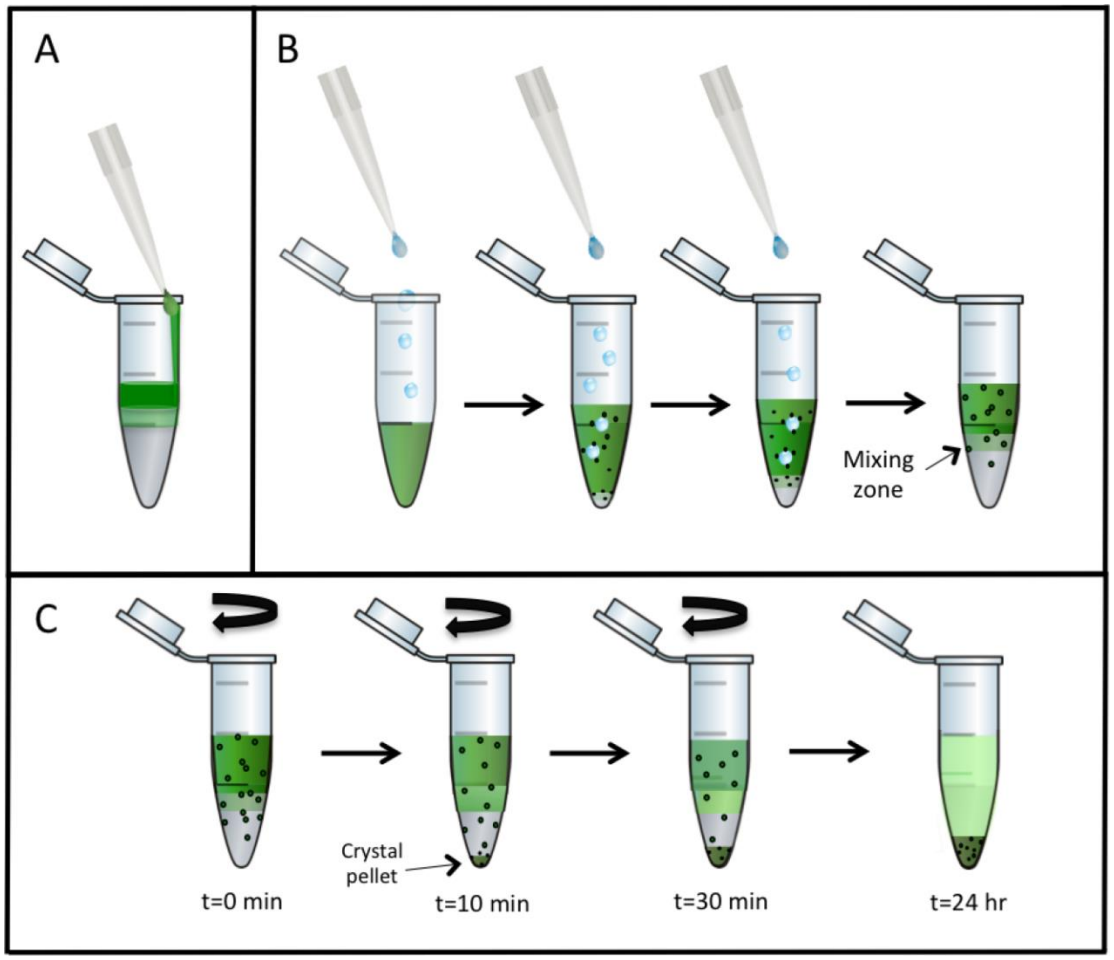


Figure 2.4. ***Schematic set-up for the crystallization experiments*** *with FID (a,b) and FID centrifugation (c). An experimental set-up (a), the protein solution is carefully layered on top of the precipitant solution, where only few crystals form at the interface. In the inverse set-up (b), the precipitant solution is added drop-wise added to the protein solution, inducing increased transient nucleation at the drop-protein interface. In (c), the experiment shown in (b) is continued by centrifugation. The nuclei formed in the protein solution are accelerated by centrifugation towards the interface zone, where they grow into nano/microcrystals, when they reach a specific size they sediment into the precipitant zone, where they stop growing. Thereby nano/microcrystals with very narrow size distribution can be achieved.*

zone. The process is fast at first and then slows down as the concentration of the protein decreases below the nucleation line. This process takes place in a time frame of 30 min to 24 h, with the majority of the crystals forming after 30 min. The crystals have a very uniform size distribution. Figure 2.5 shows a typical experiment where most of the

crystals have a radius of 500 nm. The initial screening of conditions can be performed with the same volumes as described for the FID experiments (20 µl protein plus 20 µl precipitant). The best conditions are the scaled up to a volume ratio of 50 µl protein plus 50 µl precipitant. The results obtained with this volume ration scale up very reproducible to experiments at the optimal volume ratio of 1:1. We have also twice scaled the crystallization experiments up to volumes of 6 ml protein plus 6 ml precipitant, where the experiment is performed in 15 ml Falcon Tubes.

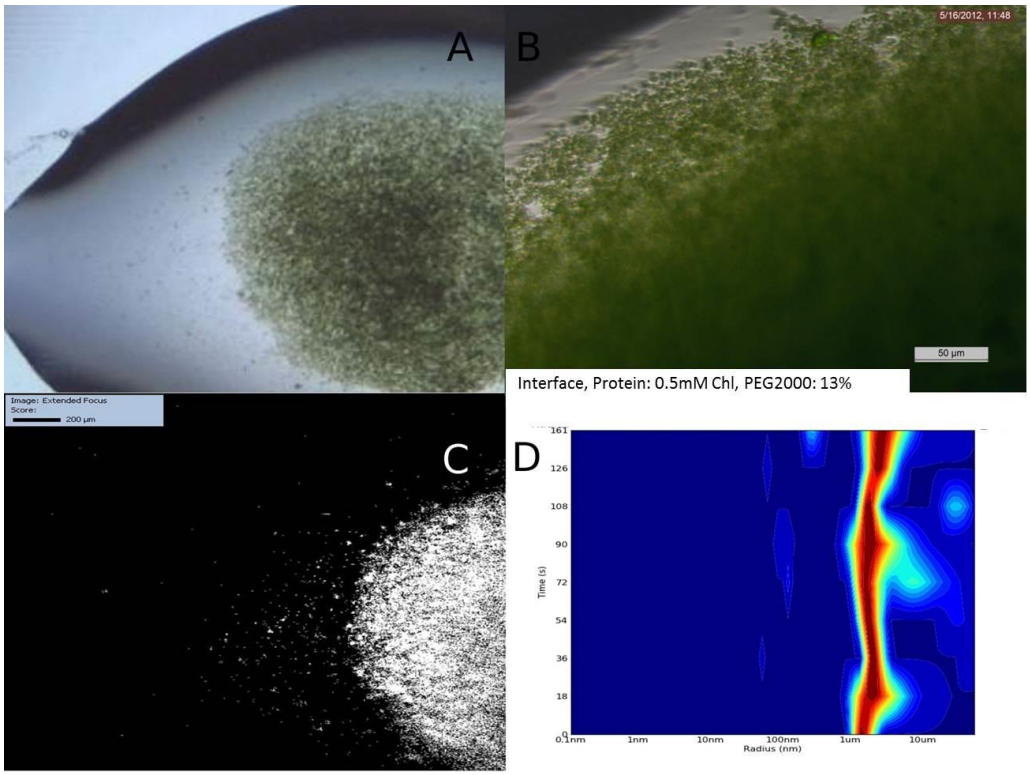


Figure 2.5. ***Free Interface Diffusion experiment using*** *0.5 mM Chl and 13% PEG2000 as starting conditions. (a,b) Images of crystals created using the FID method, crystals are significantly smaller and more uniform in size. (c)The SONICC image of the drop of (a) confirming their crystallinity.(d) The DLS histogram indicating that the majority of the crystals are 2-5 µm in radius.*

We have tested several parameters that affect the growth of the crystals by the FID centrifugation method. The size of the crystals depends on the protein concentration

and precipitant concentration and is also influenced by the viscosity of the precipitant solution. The speed of centrifugation does not have a large effect on crystal size and yield within the tested range of 200-500 rcf. The yield increases strongly with crystallization time in between 5 and 30 min. The yield is only slightly increased when centrifugation is continued to up to 24 h. Crystals of uniform size can be harvested already after 30 min. the size of the crystals grown with this technique can be varied by adjusting protein and precipitant concentration, but crystals grown with this method are consistently below 2 µm. See Figure 2.6 for examples of these crystals.

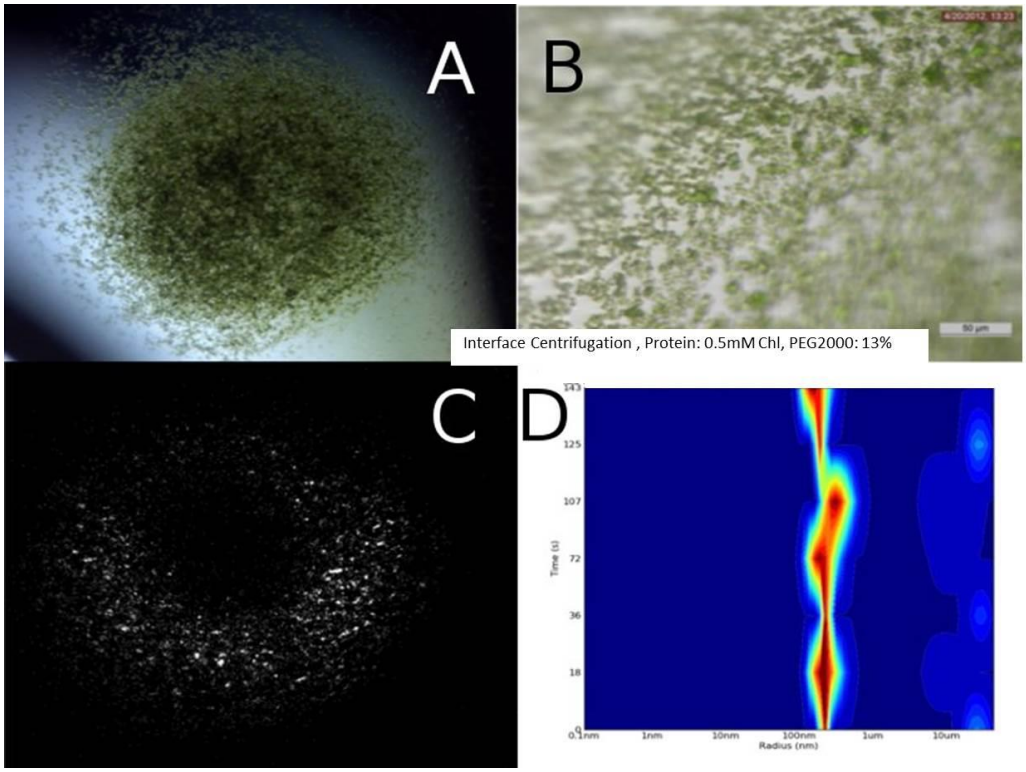


Figure 2.6. ***Free Interface Centrifugation experiment*** *using 0.5 mM chl and 13% PEG2000 as starting conditions. (a,b) Images of crystals created using the centrifugation interface method, crystals are smaller and very uniform in size and grow in approximately 30 min. (c) The SONICC image of the entire drop of (b) confirming their crystallinity.(d) DLS histogram showing that the majority of the crystals are around 500 nm in radius.*

Quenching of Crystal Growth

Ideally, crystal growth should not be quenched, as a dramatic, sudden change in conditions can lead to defects, even in nanocrystals. Data should be collected when the crystals have reached their optimal size. While this is highly desired, it is not practical.

For all methods mentioned earlier, crystal growth must still be quenched in order to avoid growing too large as the diffusion of protein will eventually allow them to continue growing. Crystal growth must be quenched as soon as they reach the desired size distribution, which requires regular monitoring of the crystal size by DLS, visualization of crystals with SONICC, and optical imaging. For example, if the crystals are grown in a 15% PEG2000 precipitant, then they should be quenched in a 20% PEG2000 solution to avoid further growth.

Crystals growth is typically quenched after they reach the desired size as monitored by optical microscopy and DLS. To quench the crystal growth, as much supernatant is removed as possible and a high concentration precipitation buffer containing 20% PEG2000 is added on top for crystal storage. These quenched crystals are best used within 2-5 days. A word of caution must be included about transport of crystals. While the crystals maintain their size distribution for days in a laboratory setting where they are stored without vibration at 10ºC, they can dramatically change their size distribution during transport, where the size distribution changes in favor of the larger crystals. The reason is that even under quenched conditions there is an equilibrium between protein in solution and in the crystals, where proteins on the surface go in solution while proteins from the solution can bind to the surface of the crystal. The surface/volume ratio is much higher for the small nanocrystals than for the larger

micrometer-sized crystals, leading to a net transfer of protein from the small to the larger crystals. This process takes very long under diffusion/convection controlled conditions when the crystal suspension is stored in a calm place with minimized vibration. However, vibration and shaking cannot be avoided during transport by air and on land, leading to the face that the size distribution changes; in the worst case scenario, most of the small crystals have completely dissolved and the remaining crystals are too large for data collection. Therefore, crystallization on the site of data collection is extremely important for the successful growth of well-ordered nano and microcrystals for SFX.

Discussion

Batch Method

While the batch technique produces large quantities of small crystals, they are suboptimal for femtosecond crystallography as there is an indication that crystallization occurs too rapidly, producing low-quality crystals which are not ideal for X-ray diffraction. Also this technique offers no size control, crystal size can range anywhere from submicrometre to 300 μm in size. This is a definite problem for SFX because crystal size must be limited to 15-20 μm when using the liquid jet (DePonte et al., 2008; Weierstall et al., 2012) as discussed above. However, there are other sample delivery techniques, including the use of the new injector developed for crystals in lipid cubic phases (Weierstall et al., 2014). This slow running jet is ideal for many SFX experiments as it requires less sample and features a much larger jet, where crystals up to 30 μm in size can be delivered. However, the slow running jet is problematic for TR-SFX experiments. Another jet type based on electro spinning has been reported by Bogan and

co-workers (Sierra et al., 2012), which also allows for the use of larger crystals, however not all crystals may survive this procedure and the hit rates reported for use if this injector are very low.

Free Interface Diffusion

This technique has several advantages over the batch technique, primarily the ability to control the size of the crystals grown. This is possibly because when the nucleation and growth occur at the interface, they are held at the interface by the high density of the underlying precipitation buffer. Once the crystals achieve a certain size, dependent upon the protein and the precipitation buffer, the crystals will sink through the precipitation buffer and pellet out. At this stage, crystal growth stops due to a lack of protein in the surrounding environment limiting the size of the crystals.

Free Interface Centrifugation

The major difference of this technique compared to the FID, comes to light when one compares results obtained by the use of a hanging bucket or fixed angle centrifuge rotor. When using a swinging bucket rotor, the crystals form a pellet to the bottom of the centrifuge where there is no protein, only high precipitant, immediately quenching the crystal growth. When using a fixed angle rotor, the crystals coat the sides of the walls, away from the centripetal force, causing a gradient in crystal sizes. Larger crystals are found towards the top of the micro centrifuge tube, and smaller cyrstals are found towards the bottom. This is most probably because the crystals at the top of the crystalline pellet are exposed to the mixing zone longer and therefore acquire more protein resulting in larger crystals.

Advantages to this technique are twofold, the first being the increased speed in which crystallization occurs. Rather than taking 24 h, a comparable yield can be achieved roughly in approximately 30 min. The second advantage is that the size of the crystals is typically much smaller; typically 0.5-2 μm in size.

Conclusion

We have presented several methods for nanocrystal growth. These methods have been tested primarily with PSII as a model protein for large unit cell membrane proteins. These techniques require a higher density precipitant to be used in the crystallization process, but many proteins are crystallized in various PEGs and other high-density media. For cases in which the protein is denser than its precipitant, this process can be reversed, with the precipitant being added to the protein. The crystallization techniques presented here have been verified on other proteins to demonstrate the generality of the method, so that we are confident that these techniques can be used for the growth of nano or microcrystals for SFX experiments.

Personal Contribution

When I first started this work, there was no condition for creating PSII microcrystals. The conditions had been relatively set for growing macrocrystals (1-5% PEG2000 with 2 mM Chl PSII), but these only gave a rough idea of where microcrystals could be produced. The first difficulty was to come up with a technique that would not only reproducibly grow small (<10 μm) crystals, but also be able to produce large quantities of crystallites of high densities.

The first method, as described above, was the batch method. This was attempted to narrow down the range of precipitant and protein concentrations that needed to be used in order to form the microcrystals. Ranges were attempted from 3-20% PEG2000 with 0.1 – 2 mM Chl. Through these first experiments I was able to determine rough concentrations at which crystallization was more prevalent than amorphous precipitate. These were determined to be between 11 – 19% PEG2000 with 0.25 – 1% mM Chl. With this method, though, the crystals were very irregular and had a huge size range. As stated repeatedly above, this does not fit the requirements for SFX; thus, a different technique was required.

The necessity for uniform crystal size required us to think about how to limit the crystals during the growth phase so that they stopped growing once the crystals reached a certain size. The free interface diffusion technique was conceived as a way to stop the growth by quenching it once a crystal reached a certain size and was able to overcome the density of the PEG2000 solution the crystals were grown in. However, because this is a different technique, it required new optimized conditions. The optimal conditions for this were determined to be between 12 – 14% PEG2000 at a concentration of 0.4 – 0.6 mM Chl. At these concentrations, crystals could reliably be made <5 μm. The secondary advantage of this technique, which I didn't anticipate but was very welcome, was that the crystallization occurred within 24 – 48 hours. This allowed for crystallization onsite at the LCLS with fresh crystals. Unfortunately, anticipating 48 hours in advance how much protein you would need can be rather difficult, and due to the involved process of purification it was our desire to waste as little protein as possible during these beamtimes.

This frugality led to the final technique for crystallization of PSII that I developed, which was the free interface diffusion centrifugation method described above. This originally was a method that I thought would only help increase the yield of smaller crystals (<2 μm) by helping the crystals to overcome the density of the PEG2000 solution. However, during the first experiment I noticed something strange and that was that after a mere 30 minutes of spinning I already had a crystal pellet the same size as what I would have expected after 48 hours of the normal FID. These crystals all turned out to be very uniform and significantly smaller than the normal. Unfortunately, due to several reasons, I was never able to extensively test these crystals at a beamtime, but the small amount of data that was collected on them suggests that they diffract to about the same quality and possibly slightly better than a normal FID crystallization. More optimization was required, however, to determine the optimal time for crystal growth, the optimal speed for centrifugation, and which type of rotor to use.

The first question I had was does rotor type make a difference. When using a hanging bucket rotor, the crystalline pellet was all formed at the bottom and the crystals were uniform in size, however when using a fixed angle I found that at lower speeds the crystalline pellet is spread up the wall of the tube with crystals getting progressively larger towards the top where they would have been exposed to more protein for longer. With the help of Chelsie Conrad and Shatabdi Roy-Chowdhury, time courses and speed variations were performed from 30 minutes to 24 hours of spinning, as well as low as 100 rcf to as high as 500 rcf. The optimal conditions were found to be around 2 hours of spin time at around 300 rcf.

While I developed these techniques specifically for PSII crystallization, we have since tested them, successfully, with many other proteins and gotten microcrystals that are usable for SFX. The one major factor necessary for this technique to work is that either the protein or the precipitant solution must have a higher density, otherwise mixing occurs too rapidly and crystallization is not as controllable.

Chapter 3

Serial Time-Resolved Crystallography of Photosystem II Using a Femtosecond X-ray Laser

Christopher Kupitz[1#], Shibom Basu[1#], Ingo Grotjohann[1], Raimund Fromme[1], Nadia A. Zatsepin[2], Kimberly N. Rendek[1], Mark Hunter[1,3], Robert L. Shoeman[4], Thomas A. White[5], Dingjie Wang[2], Daniel James[2], Jay-How Yang[1], Danielle E. Cobb[1], Brenda Reeder[1], Raymond G. Sierra[9], Haiguang Liu[2], Anton Barty[5], Andrew L. Aquila[5,7], Daniel Deponte[5,8], Richard A. Kirian[2,5], Sadia Bari[12,17], Jesse J. Bergkamp[1], Kenneth R. Beyerlein[5], Michael J. Bogan[9], Carl Caleman[5,16], Tzu-Chiao Chao[1,15], Chelsie E. Conrad[1], Katherine M. Davis[18], Holger Fleckenstein[5], Lorenzo Galli[5,6], Stefan P. Hau-Riege[3], Stephan Kassemeyer[4,12], Hartawan Laksmono[9], Mengning Liang[5], Lukas Lomb[4], Stefano Marchesini[14], Andrew V. Martin[5,10], Marc Messerschmidt[8], Despina Milathianaki[8], Karol Nass[4,5,6], Alexandra Ros[1], Shatabdi Roy-Chowdhury[1], Kevin Schmidt[2], Marvin Seibert[8,13], Jan Steinbrener[4], Francesco Stellato[5], Lifen Yan[18], Chunhong Yoon[5,7], Thomas A. Moore[1], Ana L. Moore[1], Yulia Pushkar[18], Garth J. Williams[8,], Sébastien Boutet[8], R. Bruce Doak[2], Uwe Weierstall[2], Matthias Frank[3], Henry N. Chapman[5,6,11], John C.H. Spence[2] and Petra Fromme[1]

[1]Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604 USA
[2] Department of Physics, Arizona State University, Tempe, Arizona 85287 USA
[3] Lawrence Livermore National Laboratory, Livermore, California 94550, USA
[4]Max-Planck-Institut für medizinische Forschung, Jahnstr. 29, 69120 Heidelberg, Germany
[5]Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany
[6]University of Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
[7]European XFEL GmbH, Notkestrasse 85, 22607 Hamburg Germany
[8]Linac Coherent Light Source, Stanford Linear Accelerator Center (SLAC) National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA
[9]Stanford PULSE Institute, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA
[10]Department ARC Centre of Excellence for Coherent X-ray Science, of Physics University of Melbourne, Australia
[11]Center for Ultrafast Imaging, Luruper Chaussee 149, 22761 Hamburg, Germany
[12] Max Planck Advanced Study Group, Center for Free-Electron Laser Science (CFEL), Notkestrasse 85, 22607 Hamburg, Germany

[13]Uppsala University, Olofsgatan 10B, 753 12 Uppsala, Sweden
[14]Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
[15]University of Regina, 3737 Wascana Pkwy Regina, SK S4S 0A2, Canada
[16]Department of Physics and Astronomy, Uppsala University, Regementsvagen 1, SE-752 37, Uppsala, Sweden
[17]Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany
[18]Department of Physics, Purdue University, 525 Northwestern Avenue, West Lafayette, IN 47907, USA

# These authors contributed equally to this publication

Summary

Photosynthesis converts sunlight into the energy needed to sustain life on Earth. Two large membrane protein complexes, Photosystem I and II (PSI and PSII), act in series to catalyze the light-driven reactions in photosynthesis. Photosystem II (PSII) catalyzes the light-driven water splitting process, which maintains the Earth's oxygenic-atmosphere(Renger, 2012). In this process, the Oxygen-Evolving Complex (OEC) of PSII cycles through five states, S_0 to S_4, where four electrons are sequentially extracted from the OEC in four light-driven charge-separation events. Here, we describe microsecond time resolution serial femtosecond crystallography (SFX), a recent technique(Chapman et al., 2011) applied to experiments on PSII nano/micro crystals, where structures have been determined from PSII in the dark S_1-state and after double laser excitation (putative S_3-state) at 5 and 5.5 Å resolution respectively. The results provide evidence that PSII undergoes significant conformational changes at the electron acceptor site and at the Mn_4CaO_5 core of the OEC. These include an elongation of the metal cluster, accompanied by changes in the protein environment, which could allow for binding of the second substrate water molecule between the dangler Mn and the cubane in the $S_2 \rightarrow S_3$ transition, as predicted by spectroscopic and computational

studies(Navarro et al., 2013)(Isobe et al., 2012). The work presented here shows the great potential for time-resolved SFX studies to investigate catalytic processes in biomolecules.

The first X-ray structure of PSII was determined to a resolution of 3.8 Å in 2001(Zouni et al., 2001), revealing the protein's architecture and the overall shape and location of the OEC. In 2011, Shen and coworkers achieved a breakthrough in the structural elucidation by dramatically improving crystal quality, enabling determination at 1.9 Å resolution(Umena et al., 2011). This structure showed the OEC at near atomic resolution. However, the OEC was likely affected by X-ray damage, a fundamental problem in X-ray crystallography.

The X-ray damage problem may be overcome through the use of serial femtosecond crystallography (SFX) (Chapman et al., 2011)(Boutet et al., 2012)(Redecke et al., 2013)(Spence et al., 2012), an advancement enabled by the advent of X-ray free electron lasers (XFELs). In SFX, a stream of microcrystals in their mother liquor is exposed to intense 120 Hz femtosecond XFEL pulses, thereby collecting millions of X-ray diffraction "snapshots" in a time frame of hours. Each FEL pulse is so intense that it destroys the sample; however, the exposure time is so short that diffraction is observed before destruction occurs(Barty et al., 2012).

Conventional X-ray structures correspond to a time and spatial average representation of biomolecules, leading to a "static" picture. To capture dynamic processes such as water oxidation in PSII, time-resolved X-ray data can be collected using SFX(Aquila et al., 2012)(Neutze and Moffat, 2012). Conformational changes may be observed at a time-resolution ranging from femtoseconds to microseconds by combining visible laser excitation with the SFX setup and varying time delays between

the optical pump and the X-ray probe snapshot. As partial reflections from crystals in random orientations are recorded, many snapshots must be collected for adequate sampling of the full reflections and three-dimensional reconstruction. A time-resolved pump-probe experiment was performed in 2010 using PSI-ferredoxin crystals as a model system, where changes in diffraction intensities, consistent with a light-induced electron transfer process in the PSI-ferredoxin complex and dissociation of the PSI-ferredoxin complex were seen(Aquila et al., 2012).

The catalytic reaction in PSII is a dynamic process. The oxygen evolution reaction is catalyzed by the oxygen evolving complex, where the electrons are extracted from the Oxygen-Evolving Complex (OEC) in four sequential charge separation events through the S-state cycle (Kok cycle), as shown in Fig. 3.1a (see (Renger, 2012) for a review). Recently, Kern et al. reported pump-probe simultaneous SFX diffraction and X-ray emission spectroscopy (XES)(Kern et al., 2013) to investigate the dark S_1-state and the single flash (S_2-state) of PSII. The XES data show that the electronic structure of the highly radiation sensitive Mn_4CaO_5 cluster does not change during femtosecond X-ray exposure.(Kern et al., 2013) However, the quantity and quality of X-ray diffraction data was insufficient to determine if any structural changes occurred.

We report on microsecond time-resolved SFX experiments conducted at the CXI instrument(Boutet and Williams, 2010) at the Linac Coherent Light Source (LCLS).(Emma et al., 2010) The experimental setup is shown in Figs. 3.1b and c. We developed a multiple-laser illumination scheme that progressively excites the OEC in dark-adapted PSII nano/micro crystals by two laser pulses from the dark S_1-state via the S_2-state to the double-flash putative S_3-state. Not all PSII centers progress to the next S-

state by a single saturating flash which could lead to heterogeneities. Therefore the S-state reached in the double-flash experiment is indicated as "putative S_3-state" throughout the manuscript.

The diffraction patterns collected from dark and illuminated crystals were sorted into two data sets. Using the "hit finding" program Cheetah, 71,628 PSII diffraction images were identified from the dark diffraction patterns and 63,363 were identified from the double-flash patterns. From these hits, 34,554 dark state patterns and 18,772 double-flash patterns were indexed using the *CrystFEL* software suite(White et al., 2012) (see Appendix A Table 3a,b). The data were indexed as orthorhombic, with unit-cell parameters of *a*=133Å, *b*=226Å, and *c*=307Å for the dark state, and *a*=136 Å, *b*= 228 Å, and *c*=308Å for the double-flash state. The distributions of unit cell dimensions are shown in Fig. 3.2 and Appendix A. The data clearly supports an increase in unit cell dimensions in the double-flash state, with the largest difference detected for the *a*-axis. Two factors may explain the change in unit cell constants, lower indexing rates, and a slight decrease in resolution of diffraction: crystal degradation upon laser illumination or significant structural changes upon the transition from the dark state to the double flash state, which may represent the putative $S_1 \rightarrow S_3$ transition. To distinguish between these

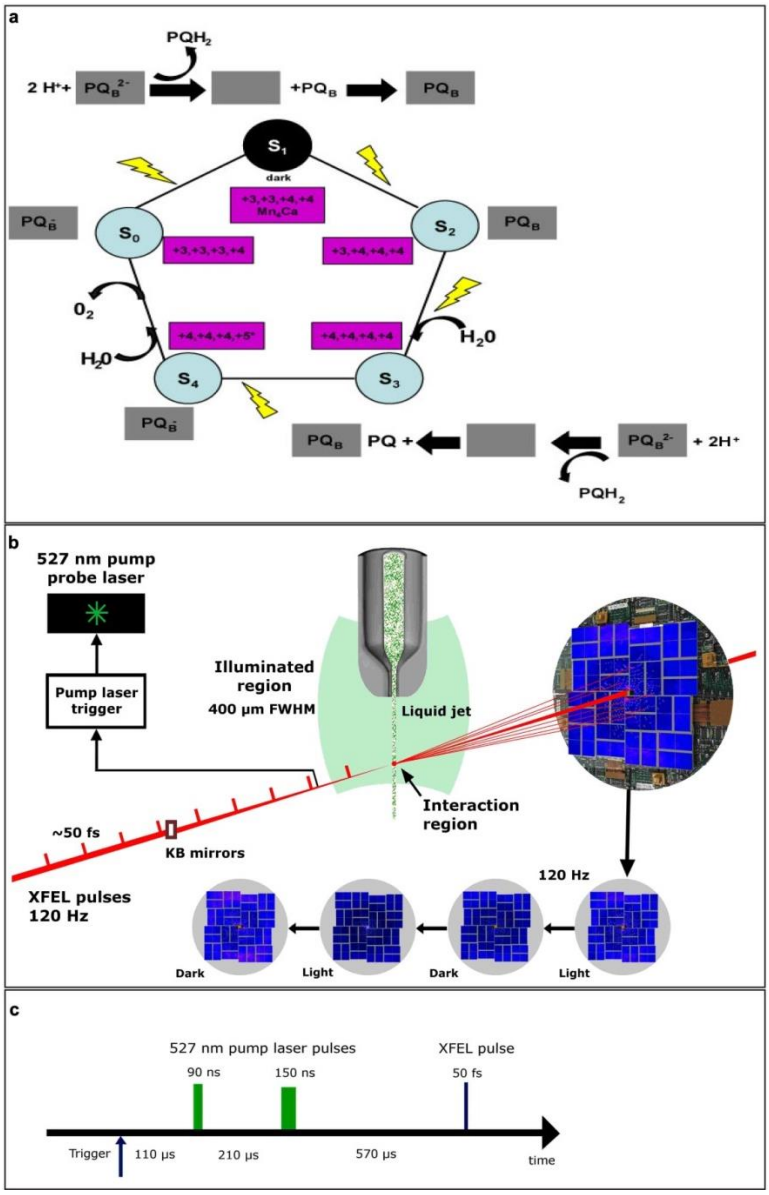


Figure 3.1. ***Experimental Schemes for the Time-Resolved Serial Femtosecond Crystallography Experiments on PSII.*** *(a) S-state scheme of the OEC depicting changes the oxidation state of the four manganese ions of the Mn_4CaO_5 cluster in the S-state cycle and indicating the reduction of the plastoquinone (PQ) to plastoquinol (PQH_2) in the Q_B site. (b) Experimental setup. The crystal-steam of PSII, was exposed to two subsequent optical laser pulses at 527 nm before interacting with the femtosecond X-ray FEL pulses. With a FEL frequency of 120 Hz and triggering of the laser at 60 Hz, X-ray diffraction patterns from crystals in the dark state and "light" double-flash state alternate. (c) Laser excitation scheme. The first 527 nm laser pulse excited the crystals 110 µs after the trigger pulse, the delay time between the first and second 527 nm laser pulse was 210 µs, with X-ray diffraction data collected 570 µs after the second laser pulse.*

two possibilities, we collected data with triple-flash excitation of the PSII crystals, where at least part of the PSII centers may have reached the putative transient S_4-state. Preliminary data evaluation of the triple-flash data set (i.e., putative S_4-state) shows similar unit cell dimensions and crystal quality as the dark S_1-state (see Fig. 3.2 and Appendix A). This suggests that conformational changes induced in PSII by the double-flash excitation (i.e., during the putative $S_1 \rightarrow S_3$ transition) are reversed after excitation with the third flash (in the putative $S_3 \rightarrow S_4$ transition). While the number of indexed patterns currently available does not yet allow for the determination of an accurate structure of the PSII after triple excitation, the data allows extraction of information on the hit rates, indexing rates and unit cell constants, showing that the unit cell constants are identical for the dark S_1 and triple-flash state. The conformational change and associated unit cell changes may be caused by dissociation of the mobile plastoquinone PQ_B from the Q_B binding pocket after double-flash excitation, when PSII may reach the S_3-state (see Fig. 3.1a). The structural changes leading to the difference in unit cell constants are likely most significant at the stromal side of PSII where the quinone bindings sites are located. To avoid structural heterogeneity at the acceptor side by partial re-occupation of the PQ_B binding site, no quinone was added to the crystals for the double pump experiments. We thereby may have "trapped" PSII in the double flash experiment in the putative S_3-state conformation with an empty Q_B binding pocket. In order to transition from $S_3 \rightarrow S_4$, an electron acceptor must replenish the empty Q_B binding site. Therefore, the plastoquinone derivative $PQ_{decyl,}$ which diffuses into the Q_B pocket,

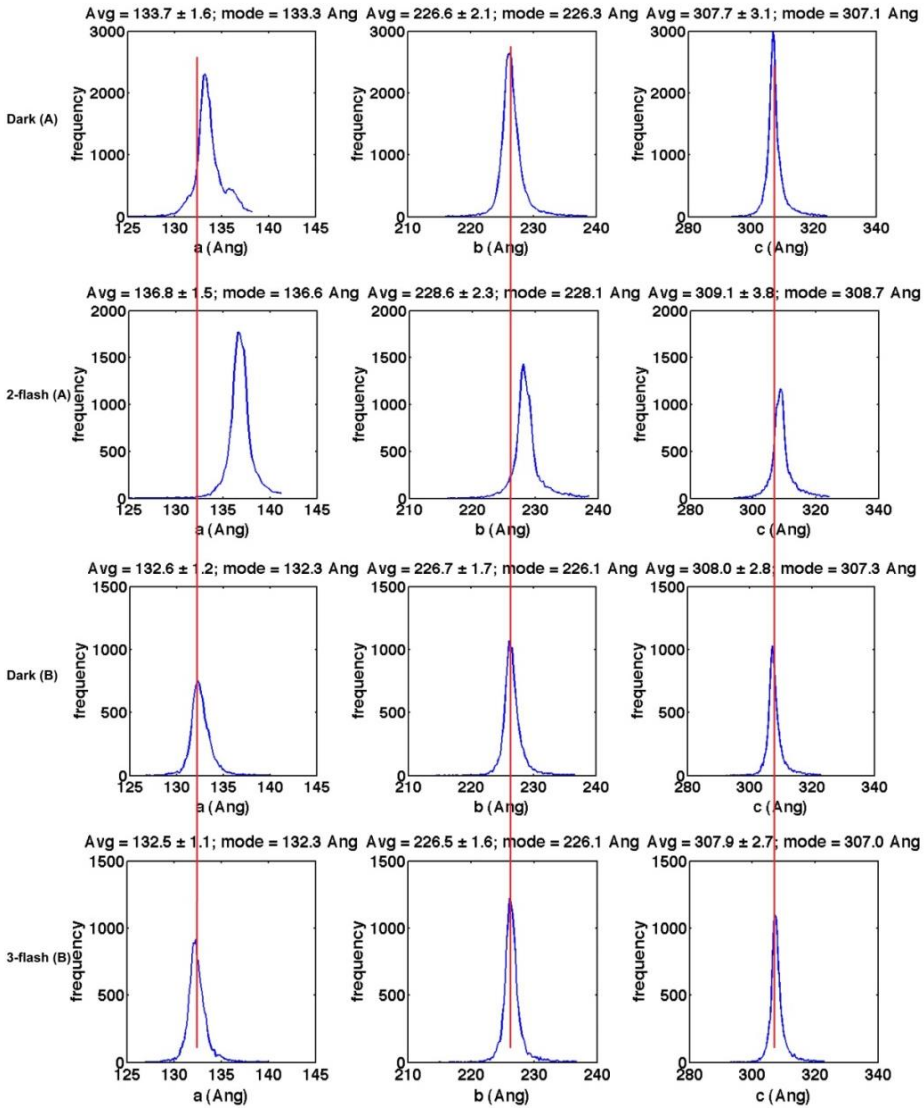


Figure 3.2. ***Distribution of PSII unit cell constants*** *of four different femtosecond crystallography data sets. (Row 1) Unit cell constants of the dark data set (S_1-state) collected at the CXI instrument in January 2012, (experiment (A)). (Row 2) Unit cell constants of the double-flash data set (putative S_3-state) collected at the CXI instrument in January 2012 (experiment (A)). (Row 3) Unit cell constants of the dark data set (S_1-state) collected at the CXI instrument in June 2012 (experiment (B)) (quinone PQ_{decyl} was added to these crystals to allow replacement of the quinone for triple excitation). (Row 4) Unit cell constants of the triple-flash data set (putative S_4-state) collected at the CXI instrument in June 2012 (experiment (B)). The comparison of unit cell constants shows that significant changes in the unit cell constants are observed after double-flash excitation of PSII. These changes are fully reversed when PSII is excited by a three laser flashes.*

was added to the crystals used for the triple-flash excitation data set. With the Q_B binding site re-occupied, the change in unit cell constants is reversed.

Diffraction data from the dark and double-flash states were evaluated to 5 Å and 5.5 Å resolution, respectively. Since each diffraction pattern represents a thin cut through reciprocal space by the Ewald sphere, only partial reflections were recorded. A high multiplicity of observations is therefore needed for each Bragg reflection to obtain accurate structure factors. The average multiplicity per reflection was 617 for the dark state data set and 383 for the double-flash data set over the whole resolution range. Data statistics tables comparing the various data sets as well as comparing the statistics between our data and other published data are shown in Appendix A.

The data were phased by molecular replacement using a truncated version of the 1.9 Å structure (pdb code 3ARC)(Umena et al., 2011). Rigid body refinement (*phenix.refine*) was performed for both the dark and double-flash structures (see methods for further details on molecular replacement and refinement). To reduce model bias, we calculated omit maps and simulated annealed maps (SA-omit maps) for the dark and double-flash data, deleting the coordinates of the Mn_4CaO_5 cluster from the model.

Figure 3.3a-c shows the arrangement of protein subunits and cofactors of Photosystem II, including the electron transport chain. The comparison of the electron density maps for the dark state (green) and the double-flash state (white) at a contour level of 1.5 σ is shown in Fig. 3.3d-f. Both maps show clear electron densities for the transmembrane helices as well as loops and cofactors. Additional electron density maps for representative structural elements of PSII are shown in Appendix B. Overall, the protein fits into the electron densities for the dark and double-flash states and matches the

high resolution structural model. However, differences appear in regions of the Mn_4CaO_5 cluster and the acceptor side, where the quinones and the non-heme iron are located. Determining the significance of these changes and their correlations is complicated by the low resolution of the data. Figure 3.3g-i shows detailed views of the loops at the acceptor side of PSII. The quinones are not visible at the current resolution of 5 Å. The maps

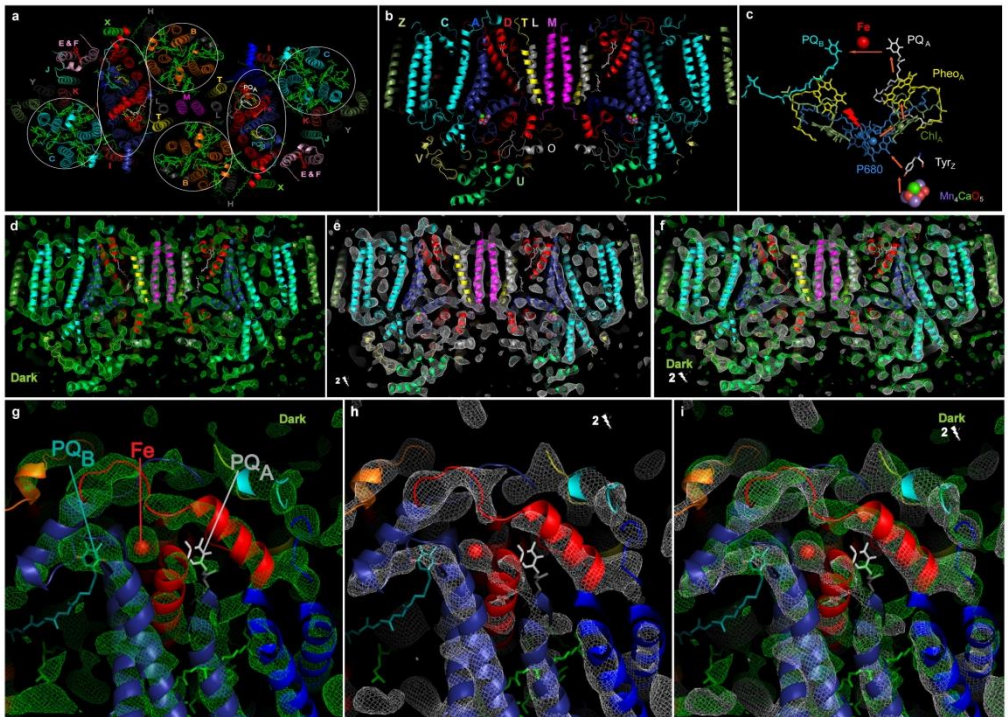


Figure 3.3. ***Overall Structure and Omit Map Electron Density of PSII.*** *(a) Transmembrane helices and cofactors in Photosystem II (stromal view). The proteins are named according to their genes and labeled with colored letters. (b) Side view of PSII at its longest axis along the membrane plane. (c) Electron transport chain of PSII (P680 (blue), accessory chlorophylls (smudge-green), Pheophytins (yellow) and plastoquinones PQ_A (white) and PQ_B (cyan)); atoms of the OEC are depicted as spheres (Mn purple, Ca green, O light red). Omit map electron densities (view as in b) at 1.5 σ for the (d) dark state (S_1) (green) (e) double-flash state (putative S_3-state) (white) and (f) overlay of the two omit maps. Omit maps (1.5 σ) of the Electron acceptor side of Photosystem II for the (g) dark (S_1) (green) (h) double-flash (putative S_3-state) (white) and (i) overlay of the two omit maps. Note that changes include a shift of the electron density of the non-heme iron.*

indicate differences between the electron densities of the dark and double-flash states in the loop regions and also in the position of the non-heme iron that is coordinated by the loops.

We now focus on the structure in the undamaged dark S_1-state of the metal cluster in the OEC and the potential light-induced structural changes that may occur during the S-state transition. Appendix B figure B-4 shows the SA-omit map of the OEC in the dark S_1-state for the Mn cluster in PSII with the 1.9 Å X-ray structure of Umena et al.(Umena et al., 2011) Interestingly, the electron-density map of the 'dangler' Mn atom from the 1.9 Å structure is located outside the dark S_1-state electron density, a feature also visible in the electron density map of Kern et al.(Kern et al., 2013). These structural observations are consistent with spectroscopic results, which indicate that the distance between the dangler Mn and the Mn_3O_xCa distorted cubane is indeed shorter in the dark S_1-state than in the 1.9 Å structure based on the synchrotron data collection, which might be influenced by X-ray induced reduction of the Mn ions in the metal cluster(Luber et al., 2011)'(Davis et al., 2013). This shorter distance is in agreement with density function theory (DFT) studies(Luber et al., 2011)'(Ames et al., 2011)'(Isobe et al., 2012) based on the 1.9 Å structure of PSII(Umena et al., 2011), however the current resolution limit of 5 Å does not allow a quantitative assessment.

The mechanism of water splitting is intensely debated and many models have been proposed. The recent 1.9 Å X-ray structure(Umena et al., 2011) formed the basis for more detailed theoretical studies of the process, yet the proposed mechanisms differ(Ames et al., 2011; Isobe et al., 2012)' (Rivalta et al., 2012; Siegbahn, 2012). Based on our TR-SFX structural data, we looked for differences between the electron-density

maps of the OEC, derived from the dark and the double-flash data-sets. Figure 3.4a and b show the SA-omit maps calculated for dark (blue) and double-flash state (yellow) and compared with the model of the metal cluster from the 1.9 Å structure(Umena et al., 2011) (Fig. 3.4c). The Mn_4CaO_5 cluster was omitted from the model for the calculation of the SA-omit map, which includes annealing at the virtual temperature of 5000 K to minimize phase bias. The SA-omit electron densities of the dark and double-flash states differ in the shape and position, as well as in the protein environment, of the Mn_4CaO_5 cluster. The dark state simulated-annealed (SA)-omit electron density for the OEC protein environment matches the model of the 1.9 Å structure (Umena et al., 2011), while the SA-omit map of the double-flash state differs significantly. Any interpretation of changes in the protein environment of the OEC is highly speculative at a resolution of 5 Å and heterogeneities in the S-state transitions. However, the SA-omit map of the double-flash state is suggestive of conformational changes which may indicate of a movement of the CD loop (including the ligand D170) away from the cluster. If confirmed at higher resolution, this could explain mutagenesis studies that questioned D170 as a ligand in the higher S-states (Debus et al., 2005). Furthermore, in the double flash state, the electron density of the metal cluster extends and shows a new connection to the AB loop at site where D61 is located. While D61 only serves as a second sphere ligand in the 1.9 Å crystal structure[6] mutagenesis studies indicated an important role in the water oxidation process as the S_2 to S_3 transition is blocked in D61 mutants.

The change in the electron-density of the OEC is suggestive of an increase of the distance between the cubane and the ‘dangler’ Mn and distortion in the cubane in the double-flash state. The observed electron densities (Figs. 3.4a and b) of the dark state and

double flash state are consistent with conformational changes predicted in a recent DFT study of the S_3-state by Isobe et al, shown in Fig. 3.4d. The increased distance between

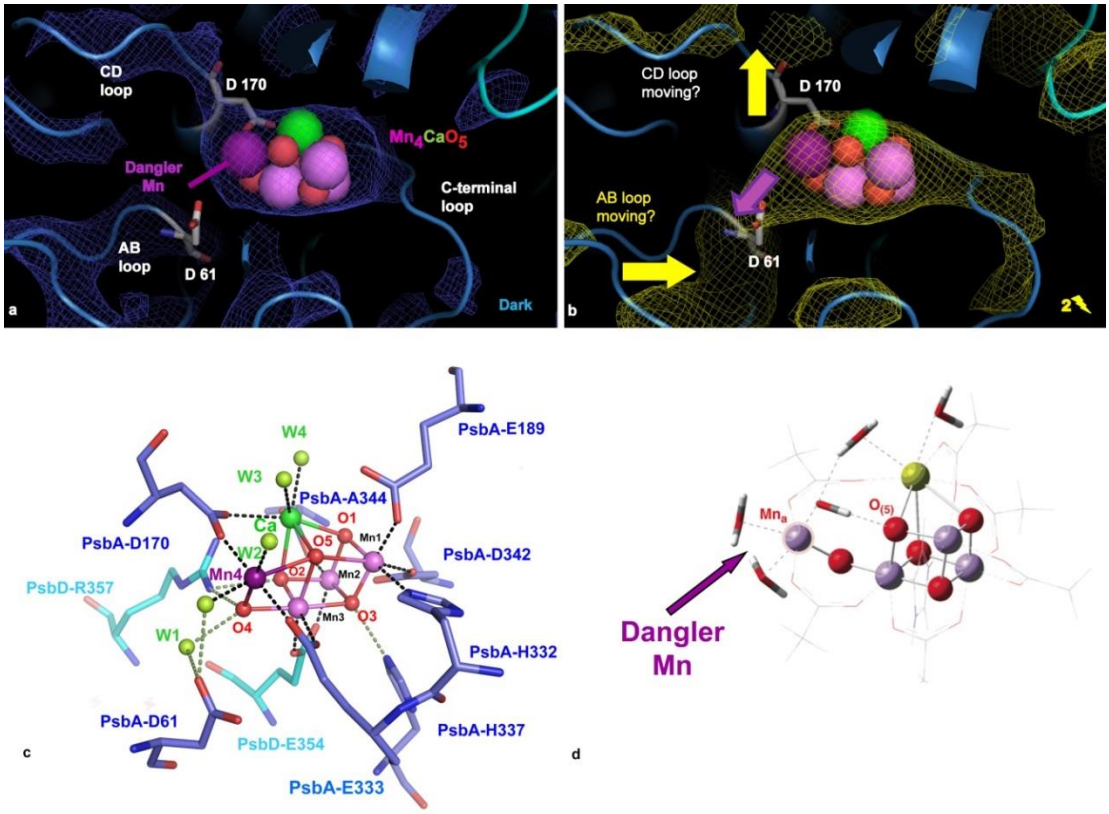


Figure 3.4. ***The Oxygen Evolving Cluster****: Simulated annealed omit maps at 1.5 σ for dark and double-flash states of the Mn_4CaO_5 cluster of PSII for the (a) dark S_1-state (blue) and (b) double-flash, putative S_3-state with the 1.9 Å structural model (3ARC) from Umena et al.(Umena et al., 2011). Mn in the distorted Mn_3O_xCa cubane (Mn-1 to Mn-3) (light-pink), dangler manganese (Mn-4) (violet), calcium (green) and oxygen (red). (c) 1.9 Å crystal structure of the Mn_4CaO_5 cluster with ligands from(Umena et al., 2011) (pdb 3ARC). (d) proposed model of the S_3-state based on DFT calculations by Isobe et al (Isobe et al., 2012) (adapted with permission.) Larger diversions in the SA-omit map of the double-flash (putative S_3-state) include potential movement of the loop connecting transmembrane helices C/D (CD loop) with D170 and AB loop (with D61), and an increase of the distance between the dangler Mn and the Mn_3O_xCa cubane (violet arrow).*

the cubane and 'dangler' Mn could allow the second "substrate" water molecule to bind between the Mn_3OxCa cubane and the dangler Mn in S_2 to S_3-state transition. It was shown by EXAFS study that the Mn-Ca^{2+} distances in the Mn_3O_xCa cubane shrink in S_3-

state(Pushkar et al., 2008). While the Jahn-Teller effect extends the distances between metals in the lower S-states of the OEC (Mn oxidation states +II, +III and +IV), a shrinking of the Mn_3O_xCa cubane is predicted in the S_3-state when all 4 Mn in the OEC have reached the oxidation state +IV. The comparison of the electron density in the dark and the double-flash states may indeed be suggestive of an overall decrease in the dimension of the Mn_3O_xCa cubane in the double-flash state, which are in good agreement with the proposed S_3- state EXAFS (Extended X-ray Absorption Fine Structure) and XES models(Dau et al., 2012).

The consistency of spectroscopy and DFT studies with our observations may provide preliminary indication that a significant fraction of the OEC centers in our crystals have reached the S_3-state in the double flash experiment. In light of new results on theoretical modeling of the OEC*(Luber et al., 2011),(Siegbahn, 2012),(Isobe et al., 2012),(Navarro et al., 2013),(Cox et al., 2013),(Rivalta et al., 2012)* we further examined the SA-omit maps in the dark and double-flash states for differences in the metal cluster that can be detected even at low resolution and discuss the results here in light of recent computational and spectroscopic studies on the metal cluster . The changes in the density of the Mn_4CaO_5 metal cluster are suggestive of an increase of the distance between the cubane and the ‘dangler’ Mn and a distortion of the cubane in the S_3-state. The observed electron densities are compared in Fig. 3.4a and b, with the recent theoretical studies of Isobe and coworkers*(Isobe et al., 2012),* shown in Fig. 3.4d, who predicted a "breakage" of the dangler Mn from the cubane cluster in the S_3-state. Additionally, EXAFS data constrains the extent of the movement of the dangler Mn relative to the cubane(Grundmeier and Dau, 2012),(Yano and Yachandra, 2008). The increase in

distance could allow for the binding of the second substrate water molecule between the dangler Mn and the Mn_3CaO_x cubane. The presence of a substrate water molecule between the dangler Mn and the distorted cubane in the higher S-states, has also been predicted to be essential for the catalytic mechanism in a recent DFT model of the full catalytic S-state cycle, including modeling of the substrate water exchange(Siegbahn, 2012),(Siegbahn, 2013).

In addition to the elongation, the overall dimensions of the Mn_4CaO_5 cluster appear to condense in the double-flash data set that may represent the putative S_3-state. This may include shrinking of the distance between the Ca^{2+} and the 3 Mn in the distorted cubane. EXAFS studies on PSII, where the Ca was substituted with Sr, showed significant changes in Mn-Mn or Mn-Sr distances in the S_3-state*(Pushkar et al., 2008)*, which were interpreted to indicate the distance between Mn and Ca would shrink in the S_3-state. Our experimental findings suggest a shrinking of the Mn_4CaO_5 cluster in double-flash state, which supports the hypothesis of a condensation of the Mn_3O_xCa cubane part of the Mn_4CaO_5 cluster in S_3*(Isobe et al., 2012)*. Models of Mn-oxygen cubane compounds show an increased distance between the Mn and O atoms in the cubane at lower oxidation states (+2 and +3) due to the Jahn-Teller (JT) effect*(Rivalta et al., 2012),(Yamaguchi et al., 2013)*. Distances derived from a recently published model Mn-O and Mn_3O_xCa cubane structures*(Kanady et al., 2011)* indicate that Mn-O distances depend on the oxidation states of the Mn-ions: the average Mn^{+2}-O distance is 2.2 Å, the average Mn^{+3}-O distance is 2.0 Å and the average Mn^{4+}-O distance is 1.8 Å. Two models have been proposed on the basis of X-ray absorption and emission spectroscopy, one describes the S_3-state as Mn (+3 +4 +4 +4) and the other proposes Mn (+4 +4 +4

+4)*(Dau et al., 2012),(Yano and Yachandra, 2007).* In the model where all Mn ions have reached the Mn^{+4} oxidation state, a significant shrinking of the dimension of the cluster is expected due to the lack of the JT distortion with the average Mn-O distance being reduced to 1.8 Å*(Yamaguchi et al., 2013).* The shrinking of the overall dimensions of the metal cluster, supported by our maps of the double-flash state, appears to be in agreement with the studies on model compounds. This indicates that the JT distortion diminishes in the putative S_3-state during progression of the S-states cycle when all Mn reach their +4 oxidation states*(Kanady et al., 2011).*

The SA-omit maps of the dark (S_1) and the double-flash (putative S_3) states may be also indicative of changes in the protein environment of the Mn_4CaO_5 cluster. While the electron density map in the dark S_1-state overall follows the protein backbone of the 1.9 Å structure, larger perturbations of the protein environment of the cluster are visible in the double-flash state. The double-flash state electron density map may suggest a movement of the loop which connects the transmembrane helices C and D at the lumenal site (the CD loop, including D170) away from the metal cluster and a movement of the AB loop (connecting the transmembrane helices A and B) into closer vicinity to the cluster, which may allow D61 to be become part of the ligand sphere of the metal cluster. While this interpretation of changes in the protein environment of the cluster is highly speculative at the given resolution, it could explain the results of mutagenesis studies on PSII. Although the mutation of D170 (which coordinates the dangler Mn and the Ca in the 1.9 Å structure of PSII) into an alanine has no strong effects on the oxygen evolution function*(Debus et al., 2005),(Chu et al., 1995),* less than 15% of the oxygen evolution function remains in the Asp61Ala mutation*(Hundelt et al., 1998),(Dilbeck et al., 2012).*

This mutagenesis result was difficult to rationalize because D61 is found only in the second ligand sphere of the OEC in the 1.9 Å structure*(Umena et al., 2011)*. However, our SA-Omit electron density map of the metal cluster in the double-flash state shows a connection to the protein electron density in close vicinity to Asp61 (see Fig. 3.4b). This finding may provide a first indication that Asp61 may serve as a ligand to the dangler manganese in the higher S-states. While details of the conformational changes cannot be unraveled at the current resolution of 5 Å, the comparison of the dark and double-flash state SA omit maps provide an indication that the protein ligand sphere of the Mn_4CaO_5 cluster may undergo significant changes when the OEC reaches the double-flash (putative S_3) state.

Our time-resolved SFX study captures the image of PSII after it has been excited by two saturating flashes and provides experimental evidence for structural changes occurring in the putative S_3-state of the OEC, accompanied by structural changes at the acceptor side of PSII. As the resolution is limited to 5 Å, the interpretation of the changes observed is preliminary. This work is a proof-of-principle that time-resolved SFX can unravel conformational changes at a moderate resolution and may lay the foundation for unraveling of the high resolution structures of PSII in all stages of the water oxidation process in the future. To unlock the secrets of the water-splitting mechanism by TR-SFX at atomic detail, the resolution must be further improved and structures must be determined from all the S-states with multiple time delays.

Discussion of Methods

Isolation and Crystallization of Photosystem II

Photosystem II (PSII) was isolated from Thermosynechococcus elongatus as described in*(Zouni et al., 2000)* with the following modifications. The samples were frozen after the ion exchange chromatography step and batch precipitation/crystallization of PSII was performed four times in decreasing concentrations of PEG 2000, the last purification step was performed at LCLS directly prior to growth of microcrystals.

Standard crystallization methods, such as vapor diffusion and hanging drop, have become the dominant techniques for the growth of protein crystals in the past decade. These methods have been optimized for very small volumes of protein that are currently not useful for serial femtosecond crystallography (SFX). Batch crystallization has been successfully used for the growth of large PSII crystals for standard crystallography*(Zouni et al., 2000)* and crystallization in batch method can be easily scaled up for large protein volumes. Growth of very small PSII crystals (approximately 1 μm) using the batch method, at high yield, requires high super-saturation conditions to enhance nucleation. This leads to a very broad size distribution of crystals, visible crystal defects, and a coexistence of crystals and amorphous precipitate.

The growth of PSII microcrystals for data collection was performed using a free interface diffusion technique that had been adapted to a batch method for a higher yield. In this approach, nucleation is initiated at the interface between the high density precipitant solution and the lower density protein solution, containing PSII detergent micelles. The protein solution was prepared by dissolving the PSII precipitate from the 4th precipitation

step (see above) in solubilization buffer A-sol (100 mM Pipes pH 7.0, 5 mM $CaCl_2$, 10 mM tocopherol, 0.03% ß-dodecylmaltoside (DDM), glycon >99.9% purity), adjusting the concentration of chlorophyll (Chl) to 0.5 mM. The crystals were grown in batch experiments in 15 mL Falcon tubes by adding precipitation buffer to a final concentration of 100 mM Pipes pH 7.0, 5 mM $CaCl_2$, 10 mM tocopherol, and 10-17% PEG 2000 to the protein solution. The optimal PEG and protein concentration was experimentally determined for each protein batch separately, in small scale experiments, before the remainder of the protein was crystallized on site at LCLS directly prior to data collection. All precipitation steps and the crystallization were carried out in darkness to avoid pre-illumination of the crystals. The PEG precipitant solution was added to the PSII solution at a rate of approximately 20 μL per second. The slow addition of the PEG precipitant, with a higher density than the protein solution, led to a two phase system, where the precipitant solution gathered at the bottom of the tube with a small mixed zone in between the top protein layer and the bottom PEG layer. Once the crystals formed and reached a sufficient size they sedimented into the precipitant solution and formed a pellet at the bottom of the tube. As the precipitant solution did not contain protein, the crystal growth stopped (see Fig. 3.5a-e). To further ensure that crystal growth had been terminated, the supernatant was removed after 48 hours and a stabilization buffer (100 mM Pipes pH 7.0, 5 mM $CaCl_2$, 10 mM tocopherol, 20% PEG 2000) was added. This buffer also served as the running buffer for delivery of the crystals to the X-ray interaction region during the time resolved SFX experiments. The supernatant was saved and later used once crystals reached sufficient size, as it continued to crystallize at a significantly slower rate, due to the decreased protein concentration. The crystallization

progress was monitored closely by taking 1 µL samples at regular intervals to determine crystal size by dynamic light scattering (DLS) (See Fig. 3.5d). Crystals were harvested and directly used for the SFX experiments after they reached a size of around 1 µm. While DLS provides the size distribution of the particles, it cannot discriminate between amorphous and crystalline particles. The crystallinity of the samples was therefore monitored by SONICC, which detects nanocrystals as small as 100 nm*(Wampler et al., 2008)*. Fig. 3.5a-e show the crystallization method and results of crystal characterization experiments. All handling steps with the crystals were performed in dim green light to limit exposure. After growth and stabilization, crystals were stored in complete darkness. All steps thereafter were done in the dark.

The plastoquinone derivative PQ_{decyl} was not added to the crystals until the beamtime in June 2012 and therefore not a part of the double-flash (i.e., putative S_3) experiments. Jesse Bergkamp synthesized PQ_{decyl} in the lab of Dr. Ana L. and Dr. Thomas A. Moore at ASU. It contains the same head group as plastoquinone but the 15 units of the isoprene tail were replaced by an n-decyl chain to improve solubility. We have independently determined that addition of PQ_{decyl} maintains full oxygen evolving activity of PSII under continuous illumination for several minutes. A plastoquinone molecule is located in the quinone binding pocket Q_B in S_1, after two laser excitation flashes, reaching the putative S_3-state, the natural plastoquinone (PQ) becomes double reduced to PQ^{2-}, which takes up two protons and leaves the binding site as plastoquinol PQH_2. The empty binding pocket is then re-populated by $PQ_{decyl,}$ before the third laser

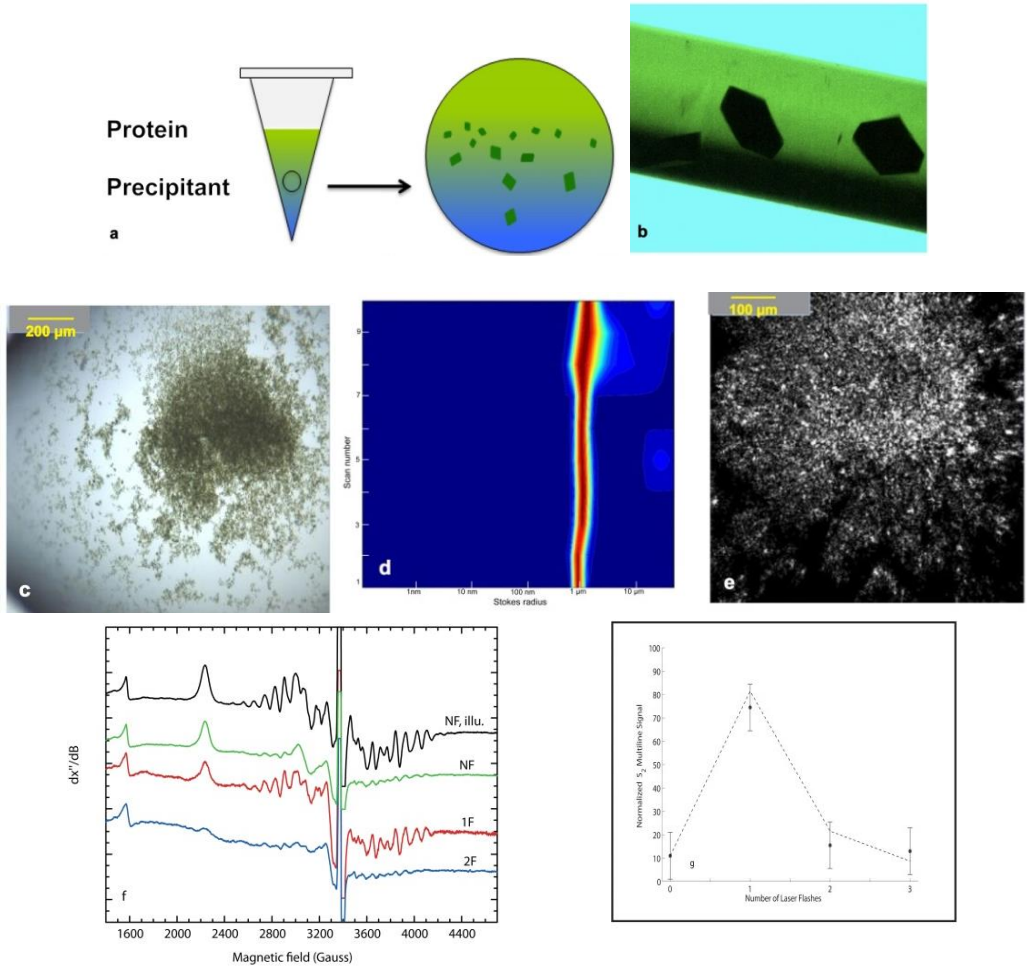


Figure 3.5: ***PSII Crystal Growth and Characterization****. (a) Scheme of free interface diffusion (FID) method for growth of PSII microcrystals (b) Large PSII crystals for data collection at synchrotron sources (c) Optical image of microcrystals of PSII grown by FID. DLS results of the crystals shown in (c) indicate an average Stokes radius of 1µm (e) SONICC image of the PSII microcrystals shown in (c) to confirm crystallinity. Panels (f-g) show the EPR analysis of S-state yield of PSII. (f) X-band EPR spectra (10 K) of PSII protein solution exposed to 0 (dark adapted sample, NF) one (1F) or two (2F) saturating laser flashes at room temperature. Samples were flash frozen after illumination. For comparison the EPR spectra of dark adapted PSII subjected to continuous illumination at 190 K (NF,illu) is shown. At low temperature, the S-state cycle stops in the S_2-state which means that this conditions corresponds to the maximal yield of multi-line signal (MLS). 3 individual samples of each type were analyzed and the same MLS intensities were consistently found for similar samples. (g) Fit of the quantified S_2-state MLS to the Kok model of the S-state transition cycle (Kok et al., 1970). Please note that the MLS yield after the 2nd and 3rd flash is nearly constant in the measurements, while the fit predicts a decline after the 3rd flash. This is expected as we have not added quinones or artificial electron acceptors to the sample, so that there is no terminal electron acceptor present after PQH_2 has left the Q_B binding site after the 2nd flash.*

flash induces the next charge separation event to reach the S_4-state (see Fig. 3.1a for the S-state scheme that also features the reduction state of the quinone in each of the S-states).

Characterization of Microcrystals by SONICC and DLS

All crystal samples were characterized via two methods, second order nonlinear imaging of chiral crystals (SONICC)*(Wampler et al., 2008)*. The SONICC and DLS experiments were performed using 24 well VDXm plates allowed for data collection of 1 µL suspension of the crystals. The reservoir was filled with 500 µL of precipitation buffer to prevent evaporation of the 1 µL hanging drop containing the crystals. The crystals were monitored by the SONICC system, (custom-equipped with a 785nm laser) at 200 mW laser power for an exposure time of 1 second. In the SONICC technique, the crystals were excited by two femtosecond infrared laser pulses of 1064 nm leading to second harmonic generation. The SONICC signal was detected at 532 nm. The dynamic light scattering experiments were performed in 1 µL hanging drops. Our DLS instrument is equipped with an infrared laser at 785 nm to avoid excitation of the pigments in PSII during the measurements. Attempts to conduct DLS measurements on PSII samples with a red laser, as used in most commercial DLS instruments, failed due to the strong absorption of red light by the chlorophylls in PSII, which strongly diminished the signal. For each sample, 10 measurements were performed with 20 seconds of data collection per measurement. The viscosity of the buffer solutions was determined experimentally by calibration with 140 nm polystyrene beads. The crystal size distribution was found to be around 1 µm (see Fig 3.5d).

EPR Characterization of the S-State Transition

Electron paramagnetic resonance (EPR) has been used extensively to determine the progression of PSII through the S-state cycle*(Han et al., 2008),(Han et al., 2012).* Using this technique, quantification of the S_2-state is determined by the multiline signal (MLS), a signature of only this state of PSII. Protein solutions, used for PSII crystallization, were cycled through the S-states by multiple single flash laser excitation (1-3 flashes) at room temperature. The samples were flash frozen directly after laser excitation in liquid nitrogen and the yield of multiline signal was interrogated by EPR at low temperature. Please note that the EPR experiments were performed under conditions that were as similar as possible but not identical to the LCLS experiments (for example the EPR data collection required freezing in glycerol, while SFX data are collected at RT without glycerol addition). As the S-state yield is an estimate, the double-flash state is indicated as "putative S_3-state" in the manuscript. Prior to illumination, glycerol was added to samples as a cryoprotectant, yielding a final concentration of ~30% by volume. This resulted in a final PSII concentration of 1.8 mg Chl/ml (2 mM). Dark adaption was performed prior to the EPR experiments, therefore the PSII samples were initially, predominately in the S_1-state. We did not attempt to get all the PSII into the S_1-state, using pre-illumination flashes in the presence of artificial electron acceptors followed by dark adaptation (as described by*(Han et al., 2012)*) as the natural mobile plastoquinone (1 Q_B per reaction center) would leave the binding site and consequently be lost in the pre-illumination phase.

For flash illumination at room temperature (20°C), a Continuum Surelite EX Nd:YAG laser was used with a second harmonic generator yielding 532 nm, 8 ns, 1 Hz,

and ~380 mJ (fluence of ~10^3 mJ/cm^2) pulses. Low-temperature X-band EPR spectra of the flash-frozen samples were recorded using a Bruker EMX X-band spectrometer equipped with a X-Band CW microwave bridge. The sample temperature was maintained at 10 K by an Air Products LTR liquid helium cryostat during collection of the EPR spectra. Spectrometer conditions were as follows: microwave frequency, 9.48 GHz; field modulation amplitude, 25 G at 100 kHz; microwave power, 31 mW.

Dark-adapted samples of both PSII solutions and crystal suspensions (frozen without illumination) contained a small percentage (typically 10%) of multiline signal. To determine the maximal possible yield of the MLS, the dark adapted, frozen PSII samples were illuminated at 190 K (dry ice/ethanol bath) for 20 minutes. Please note that the S-state cycle stops at the S_2-state at low temperature (190K)*(Han et al., 2008)*, therefore all photoactive PSII reaction centers can be brought into the S_2-state by low temperature continuous illumination. The illumination of the frozen crystal suspensions was performed in 2 minute intervals with the maximum MLS signal intensity achieved within the first 2 minutes of continuous illumination. We observed that the presence of glycerol affects the intensity of the MLS. Solutions and crystals to which no glycerol was added exhibited lower MLS intensity after continuous illuminations.

Fig. 3.5f shows the EPR spectra of PSII samples, which were excited by 1 or 2 laser flashes at room temperature, followed by flash freezing in the dark. In addition the graph also shows the control experiment where dark adapted frozen PSII was continuously illuminated at 190 K to achieve the maximal S_2 yield. A miss parameter α=9.7% was obtained by fitting the MLS intensities as a function of laser flashes, see Fig. 3.5g. Samples exposed to three flashes were also included. The data evaluation

indicates that with two flash illuminations at least 70% of the PSII reaction centers have reached the S_3-state under these conditions.

The transition rates are comparable to results of EPR studies on spinach PSII by Styring et al., who published transition rates of max 75% under conditions that were highly optimized for maximal yields of S-state transitions*(Han et al., 2012)*.

CXI Instrument Setup and Sample Delivery for Time-Resolved Femtosecond Crystallography Data Collection on PSII Crystals in the Double-Flash State

Time resolved femtosecond X-ray crystallography data were collected at the Coherent X-ray Imaging (CXI) instrument*(Boutet and Williams, 2010)* at the Linac Coherent Light Source (LCLS)*(Emma et al., 2010)* at SLAC National Accelerator Laboratory. The PSII crystals were delivered to the interaction region with the FEL beam as a suspension of crystals using the gas focusing liquid injector described in*(Weierstall et al., 2008),(Weierstall et al., 2012),(DePonte et al., 2008)*. The injection process was improved by the invention of an anti-settling device*(Lomb et al., 2012)*, which also was modified to permit temperature control of the sample. Stainless steel syringes containing the crystal suspension (pre-filtered through 10 µm stainless steel filters from IDEX) were mounted on a rotating holder, which was cooled with a Peltier element to 10°C. This setup maintained the crystals at their growth-temperature until their delivery to the FEL interaction region by the gas focusing jet. The glass capillary nozzle tips were polished to allow for visible laser excitation of the crystals in the nozzle tip. A black coating upstream of the nozzle tip prevented pre-excitation of the crystal suspension upstream of the optical laser interaction region. The gas focused liquid jet had a diameter of 4 µm at the intersection with the X-ray focal area of 2 μm^2 FWHM (full width at half maximum)

using the CXI instrument. Data were collected at the X-ray photon energy of 6.0 keV (2.05 Å) with X-ray pulse duration of approximately 50 femtosecond. The X-ray diffraction patterns were detected on a Cornell-SLAC Pixel Array Detector (CSPAD)*(Philipp et al., 2011),(Herrmann et al., 2013)*. The detector consisted of 64 panels, each 194×185 pixels tiled to span approximately 1728×1728 pixels with gaps between the tiles and approximately 19 cm along one side.

Double Laser Excitation of Photosystem II Crystals

PSII was excited by two subsequent optical pump-laser pulses from a diode-pumped, frequency-doubled Nd:YLF laser (Evolution-30, Coherent), emitting visible laser pulses at a wavelength of 527 nm. The laser was fiber-coupled from a table outside the experimental chamber, channeled into the chamber and onto the head of the liquid jet injector*(Weierstall et al., 2008; Weierstall et al., 2012)*. This wavelength provided a good compromise between transmission and absorption of light in the PSII crystals to ensure approximately homogeneous excitation throughout the crystals (the size of crystals was approximately 1 µm), as identified by DLS (See Fig.3.5d). The optical double pulse was produced by an active Q-switch with "on times" chosen such that the pulse energies of both laser pulses match. This resulted in pulse lengths of 90 nanoseconds and 150 nanoseconds for the first and second pulses, respectively this was done to maintain the total number of photons incident on the crystal same. The laser was focused to an area of approximately 400 µm in diameter with a flat top profile and aimed at the transparent tip of the nozzle, about 100 µm upstream from the X-ray interaction region. The laser beam diameter and aim point were chosen based on the desired pump-probe delays and calculations of sample flow profile and flow speed inside the capillary (average speed of

85 mm/s for 50 μm inner diameter of the nozzle) and in the jet (12 m/s for the 4 μm jet). This allowed for illumination of the crystals at the tip of the nozzle and in the liquid jet. It also ensured that crystals probed by X-rays were first exposed to both optical laser pulses for the "pumped" measurements (see Fig. 3.1b, c). The molar extinction coefficient was determined from dissolved PSII crystals at 527 nm, which was then used to calculate the absorption of PSII crystals with approximately 1 μm path length. From these parameters, it was calculated that a minimum fluence of 2.3 mJ/cm^2 (or a pulse energy of 3 μJ for a 400 μm diameter spot) was required to excite every PSII complex in a crystal of 1-2 μm diameter.

During the experiment, the laser pulse energy was monitored using a power meter placed at a 50:50 beam splitter on the laser table (50% of the energy going into the experiment). The energy of the laser pulse transmitted to the sample was calculated by the previously measured transmission of the entire fiber setup from the 50:50 beam splitter to the final lens in the injector (including fiber couplings and feed-through) and using 20% as a conservative estimate for the light transmission through the optically transparent end of the nozzle to the sample. That transmission could not be measured directly, so it was indirectly estimated. The laser pulse energy was chosen to correspond to approximately 6 μJ per pulse at the sample, i.e., three times what is required to optimally pump a 1-2 μm PSII crystal at 527 nm*(Han et al., 2012)*. The desired timing of the optical laser pulses with respect to the X-ray pulses from LCLS*(Emma et al., 2010)* were achieved by using the LCLS/CXI event generator and event reader (EVG/EVR) system that provided precise timing signals less than 1 μs before every X-ray pulse (or every other X-ray pulse) and a SRS DG645 delay generator to produce properly timed double trigger signals for the laser Q switch. The time delay between the two optical laser

pulses was set to be 210 μs corresponding to three times the time constant for the S_1 to S_2 transition of 70 μs*(Dekker et al., 1984).* The time delay between the second optical pulse and the X-ray pulse was 570 μs to allow the Oxygen-Evolving Complex (OEC) to complete the S_2 to S_3 transition (3 times the time constant of 190 μs for the S_2 to S_3 transition)*(Dekker et al., 1984).* As the electron transfer between Q_A to Q_B takes place in the 200 to 400μs time range*(de Wijn and van Gorkom, 2001),* (depending also on the organism and oxidation state of Q_B) the delay time of 780 μs between flash 1 and the X-ray pulse represents a reasonable minimal delay time aimed at following both the processes at the donor site in the OEC and the acceptor site at the Q_B binding site. The uniform change in unit-cell dimension in the double-flash experiment (putative S_3-state) and its reversion in the triple flash experiment (putative transient S_4-state) provides an independent indication for significant and uniform progression of the OEC through the S-states.

In order to monitor the delay between the pump beam and the probe beam, we separately measured the response of two photodiodes to the optical and X-ray excitation with an Acqiris digitizer. One photodiode was exposed to stray optical light on the laser table and the other to X-rays scattered from the sample jet at very high angle. This was necessary because of the disparity in the response of the diodes to each signal near the interaction region--i.e., at the required optical and X-ray fluences needed to perform the experiment, the signals could not be discriminated when using a single photodiode. The delay between the signals introduced by the measurement of the optical light on the laser table was measured at the beginning of the experiment by equalizing the signals from each beam near the X-ray interaction point. That constant was used to calculate a

calibrated delay time with sufficient precision for our experiment using the digitizer trace.

The LCLS timing signal triggered the preceding optical laser pulses for every other X-ray pulse allowing us to acquire alternating diffraction images from “dark” (ground state) and “double-flash” PSII crystals in the putative S_3-state, i.e., SFX data were collected with a frequency of the X-ray pulses of 120 Hz while the frequency of the pump laser pulses was 60 Hz (Fig. 3.1b, c). This approach minimizes other sources of error that might occur from systematic error. With this setup, alternating "dark" and double excited “illuminated" images were collected at 3600 dark images and 3600 illuminated images per minute. In addition to this alternating experimental scheme, dark run data was collected while the 527 nm pump laser was switched off (see Appendix A Table 3b for data statistics of dark only and alternate dark/light runs). Representative diffraction patterns of the S_1 and the double-flash data sets are shown in Figs. 3.6a, b.

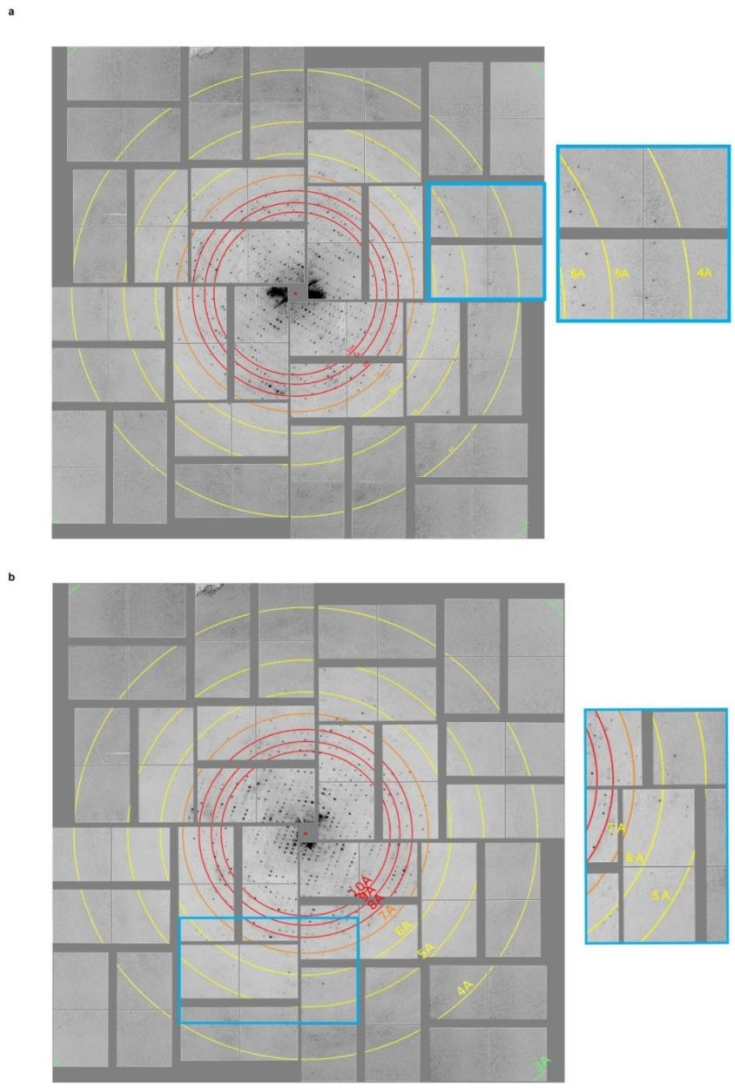


Figure 3.6. ***Background corrected diffraction pattern of a PSII microcrystal*** *from (a) the dark (S_1) data set and (b) the double-flash data set collected at the CXI instrument at LCLS. The resolution is indicated by red and yellow rings corresponding to resolution shells in Å 10, 9, 8 (red), 7 (orange), 6, 5, 4 (yellow). The right panel shows an enlarged view of the diffraction patterns (see blue box).*

Processing and Evaluation of Data with the Cheetah and CrystFEL Programs

A total of 5,528,071 raw diffraction frames were collected from Photosystem II microcrystals in January 2012. The raw diffraction data (XTC format) were pre-

processed by the *Cheetah* software package, (http://www.desy.de/~barty/cheetah, Barty et al, in preparation) and then analyzed in the software suite CrystFEL(White et al., 2012). Examples of diffraction patterns of the PSII crystals are shown in Fig. 3.6a for the dark (S_1) state and in Fig. 3.6b for the double-flash (putative S_3) state.

The first step of pre-processing involved dark current subtraction from each diffraction frame and masking of dead, hot, cold, and saturated pixels. It also included masking of the low resolution scattering from the liquid jet and the detector panel edges. Local background correction step was applied to the raw diffraction frames during their pre-processing in Cheetah (http://www.desy.de/~barty/cheetah) (publication Barty et al. in preparation). Each diffraction frame was analyzed and identified as a crystal "hit" only if it contained 25 or more peaks with the intensities of at least 400 analog to digital units. The locally background-corrected patterns collected in the alternating mode were sorted into two data sets for the dark and double-pumped hits based on signals from a photodiode and video camera in the experimental chamber. Comparison between the dark data sets from the alternating light/dark runs and complete dark runs showed no difference in unit cell constants, or indexing rates, therefore the dark data sets were merged. The final data sets contained 71,628 "hits" from the dark state data set and 63,363 "hits" from double-flash data sets. The triple-flash data sets, leading to the putative transient S_4-state of PSII (collected during a separate beamtime in June 2012), were collected in alternate runs with the frequency of 120 Hz of the LCLS X-ray pulses and 60 Hz laser excitation frequency. The same laser excitation scheme as described before was used to reach the S_4-state. The third laser excitation, reaching the S_4-state, was achieved by triggering a third laser excitation 527 nm laser pulse 570 μs after the second

laser excitation pulse. This third pulse was triggered by a second 527 nm laser installed in the CXI hutch, which excited the jet through a separate hole in the shroud. The delay time between the third laser flash and the FEL X-ray pulse was 250 µs. The data sets from the January 2012 beamtime are designated as dark (A) and double-flash (A) and the data sets from the June 2012 beamtime are designated as dark (B) and triple-flash (B) (see Appendix A Table 3a). The dark (B) and triple-flash (B) data sets contained 33,373 and 32,190 hits, respectively.

The "hits" for all data sets were passed as separate sets to the CrystFEL software suite*(White et al., 2012)* for auto indexing with MOSFLM, using the orthorhombic unit cell dimensions of PSII from Thermosynechococcus elongatus (PDB code 1FE1)*(Zouni et al., 2001)* within a tolerance limit of 6%, 5%, 5% for the reciprocal axes of a, b, c respectively for the S_1-state. Similarly, the tolerance limit of 8%, 5%, 5% were used for the reciprocal axes of a, b, c respectively for the double-flash and triple-flash states. After indexing, the 4 data sets were handled separately and the "indexing rates" (fraction of "hits" which could be successfully indexed) were 48% for dark (A), 29% for double-flash (A), 35% for dark (B), and 39% for triple-flash (B). Appendix A Table 3a, b shows the indexing statistics as well as the unit cell constants for all 4 data sets. The unit cell is orthorhombic and shows very significant changes in the unit cell dimensions between the dark S_1 and double-flash data sets (A) state (see Appendix A Table A-3 and Fig. 3.2). The most pronounced change is in the dimension of the a axis which increases by 3.3 Å. This change in unit cell constants is accompanied by the slight decrease in diffraction quality (5 vs. 5.5 Å resolution) (see Appendix A Tables A-2, A-3,) and significant lowering of indexing rates. This change in the unit cell dimension is fully reversed to the

unit cell constants observed in the dark S_1-state when PSII crystals are excited by three laser flashes, eventually reaching the putative transient S_4-state. Furthermore, the indexing rate for the dark (B) and triple-flash (B) data sets are comparable, with 35% and 39%d, respectively.

Multiplicity in the measurements is very important due to the partial nature of all reflections, the variation of the flux in each FEL pulse and the fact that each diffraction pattern is collected from a different crystal. The triple-flash data set, with 12,500 diffraction patterns, is sufficient to accurately determine the unit cell constants. However, this is border line for the determination of accurate structure factors and therefore, further data evaluation has been limited to the dark (S_1) and double-flash data sets.

The dark (S_1) and double-flash (putative S_3) data sets that were used for structure factor determination consist of 34,554 and 18,772 indexed patterns, respectively. Our data sets were merged separately in three dimensions and the structure factors were extracted separately from dark and double-flash data sets of PSII protein using the Monte Carlo method(Kirian et al., 2010), which integrates the snapshots partial reflections from randomly oriented crystals of varying size and shape (see Extended Data videos for a graphical representation of the structure factor amplitudes of the both data sets). Average multiplicities are 684 and 373 in all resolution shells from 19.20 Å to 4.03 Å of the dark (S_1) and double-flash states, respectively (see Appendix A Tables A-2a, b).

A comparison of the data statistics of our work with that of Kern et al. 2013(Kern et al., 2013) is shown in Appendix A Table A-3c. Our data sets show significantly higher multiplicity of the data and better correlation coefficients ($CC_{1/2}$) when compared to Kern et al., which are indicative of the quality of the merged reflections.

The internal consistency of the SFX data is expressed as R_{split}(White et al., 2012). For the calculation of R_{split} the sets of images of a data set are split into two random halves, and the structure factors are calculated separately for each half. The difference between the amplitudes of each of the structure factors of (hkl) plane between the two half data sets is used to estimate the convergence of the full dataset as given by R_{split}. i.e., an R_{split} of 0.22 means an estimated 22% error of the structure factors of the full data set to the "true" dataset. (The two half datasets would have agreed to 0.22 * sqrt(2).)

R_{split} was calculated for the dark and double-pumped data sets separately. As expected, with an increasing number of indexed patterns, R_{split} decreases as the data set reaches higher multiplicity and completeness (see Fig. 3.7). The average R_{split} values over all resolution shells were 0.07 for the dark and 0.09 for the double-pumped data sets.

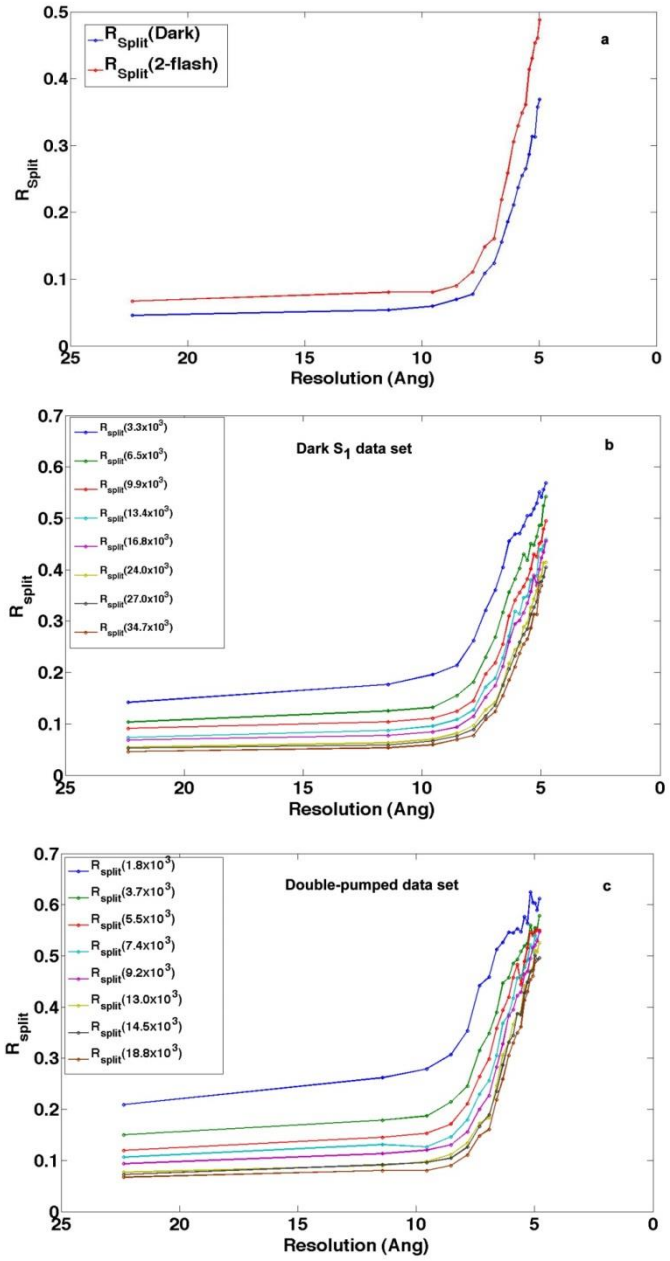


Figure 3.7. ***R_{split} as a function of resolution*** *bins and number of indexed patterns. (a) R_{split} as a function of the resolution shell (in total 20 bins) for dark state data (blue) and double-pumped state data (red). (b) R_{split} as a function of resolution bins for dark S_1-state, R_{split} decreases indicating better data quality with increase in number of indexed patterns from 3,300 to 34,000 images. (c) R_{split} as a function of resolution bins for the dark and double-flash states, the R_{split} decreases indicating better quality with increase in number of indexed patterns from 1,800 to 18,800 images.*

Structure Determination

The data for the dark (S_1) and double-flash (i.e., putative S_3) state of PSII were handled as two completely separate data sets for the whole data analysis process. After processing the raw serial femtosecond X-ray crystallographic data, the final structure factors from dark and double-pumped data sets were passed to the CCP4 Software package*(McCoy, 2007)*.

Molecular Replacement

Molecular Replacement(Read, 2003) was carried out by the PhaserMR program, which is a part of the CCP4 suite(McCoy, 2007) using the PSII X-ray structure at 1.9 Å resolution of Umena and coworkers (PDBID: 3ARC)(Umena et al., 2011) as the search model, which was modified by removing waters, detergents, lipids, and alternative conformers of the amino acids. We used monomer 1 of the PSII dimer (where the monomer 1 subunits are labeled with capital letters and small letters are used for monomer 2 in 3ARC model) as search model for molecular replacement to solve the phase problem by using the program phaser (version 2.5.3). After we found monomer 1, we repeated the search for monomer 2.

Refinement

Both structures were refined by *phenix.refine*(Afonine et al., 2012) using rigid body refinement, where each C-alpha chain of each protein subunit was considered as a rigid body. The cofactors were also considered as rigid bodies. During the rigid body refinement, we considered only translational refinement and not rotational refinement of the rigid entities. So, the RMS bond angle was not refined. We used the original B factors

from the 3ARC model because B-factor refinement is not useful at the given resolution of 5.0 Å. After three refinement cycles, R-factors for the dark state at 5.0 Å of R_{work} of 0.260 (R_{free} = 0.262) were observed. The R-factors for the double-flash state at 5.5 Å, are R_{work} of 0.280 (R_{free} = 0.290). Refinement statistics are shown in Appendix A Table A-1. All figures displaying structures were made using PyMOL (DeLano Scientific; http://www.pymol.org).

Calculation of Electron Density Maps

The electron density maps for dark and double-flash data sets were generated using the FFT program from the CCP4 suite. The omit maps, defined by Bhat, were calculated using "Omit" (CCP4 suite)(Bhat, 1988),(Vellieux and Dijkstra, 1997) where the OEC was removed from the MR solution. These ($2F_o - F_c$) omit maps were calculated using experimental data and the MR model excluding the OEC before applying refinement in order to avoid model-bias. For superimposition of these two omit maps from dark and double-flash states respectively, the following steps were carried out. First, using the omit maps as inputs, new sets of coordinate files (pdb files) were generated from the MR solutions of dark and double-flash states separately, so that each of the two omit maps fits the model, using the Molrep program in CCP4 suite. Second, the modified coordinate files for the dark and double-flash states, outputs of Molrep program, were opened, and the superimposed coordinates were saved in the coot program(Emsley et al., 2010). The double-flash state coordinate file was considered as the moving object and the dark state coordinate file as the fixed object. As a result, coot provided Euler angles and translational coordinates (x, y, z values) for this superposition. Third, using these Euler angles and translational coordinates as rotational operator with opposite sign, the double-

flash state omit map was rotated using the MAPMASK program in the CCP4 suite. Because the unit cell constants of the dark and double-flash electron density maps differ, they are in different frames of inertia, which we have to take into account for the overlaying process. The rotated double-flash state omit map (output of MAPMASK program) was moved over the superimposed coordinate file. The same procedure was applied to the dark state omit map aligning with the superimposed double-flash coordinate file using MAPMASK program in the CCP4 suite. Examples of the electron density maps are shown in Appendix B Figs.1-3.

Calculation of Simulated Annealed Omit Maps

The solutions from the molecular replacement for the PSII dimer for the dark and double-flash states were used for the calculation of the Simulated Annealed (SA) omit maps(Hodel et al., 1992). For each of the PSII coordinate files of the MR solutions, the OEC was removed and then the resulting PSII coordinate file was used for calculating the SA-omit map with a starting temperature of 5000 K using the "AutoBuild create omit map" program from the Phenix suite (version 1301 dev)(Terwilliger et al., 2008). The SA omit maps of the OEC in the dark state (S_1) and the double-flash (putative S_3) state are shown in Figs. 3.4a, b and also see Appendix B Fig. 4 for the dark state SA omit map from a different view-point.

Personal Contribution

The work detailed above was a massive international collaborative effort. In this effort I was the primary person in charge of sample preparation and characterization.This meant I was in charge of performing the PSII purifications that would be needed for the

beamtime, determining sample crystallization conditions, and crystallizing and characterizing the sample during the beamtime.

PSII purifications are an extremely long process. Thus, in order to have enough protein, more than 30 preps were needed in total, when counting the preps needed to test crystallization conditions. The preps were split between night and day shift. I primarily handled day shift starting from the whole cells through to the first HPLC runs as detailed above. Afterwards Dr. Ingo Grotjohann continued the final HPLC runs through to the freezing of the protein.

The sample crystallization conditions are examined in much more detail in Chapter 2, and therefore will not be discussed here.

During the beamtime, the sample is made fresh in an attempt to get optimal protein crystals. For this, multiple measurements were taken of each sample. As stated above, microscope, DLS, and SONICC images were taken at multiple times in order to monitor growth in an attempt to get the optimal size of crystals. I was primarily in charge of working shifts with other sample preparation team members in order to have the amount of crystals necessary each day and ensure that they were all properly characterized. I kept track of how much was used, where all samples were, and at what time points all samples in preparation were.

This was an important work because it was the first proof of time-resolved femtosecond nanocrystallography where structural changes are actually visible. Despite the lower resolution, it is very promising that changes can be seen at all, even if their exact magnitude can’t be yet determined. This is also very important work based upon its content. Oxygen evolution, as discussed at the beginning of Chapter 3, is an incredibly

important process which supports life on the entire earth. The unraveling of how PSII creates oxygen from water would be a huge step, especially in the field of bioenergetics. Understanding how it works might allow us to replicate the process thus allowing us access to new renewable resources. Thus, this is an extremely exciting and important discovery only made possible through the use of SFX and the collaboration of a great many people.

Acknowledgements

Experiments were carried out at the Linac Coherent Light Source (LCLS), a national user facility operated by Stanford University on behalf of the US Department of Energy (DOE), Office of Basic Energy Sciences (OBES). This work was supported by the following agencies, the Center for Bio-Inspired Solar Fuel Production, an Energy Frontier Research Center funded by the DOE, Office of Basic Energy Sciences (award DE-SC0001016), the National Institute of Health (award 1R01GM095583), the US National Science Foundation (award MCB-1021557 and MCB-1120997), the DFG Clusters of Excellence "Inflammation at Interfaces" (EXC 306) and the "Center for Ultrafast Imaging"; the Deutsche Forschungsgemeinschaft (DFG); the Max Planck Society, the Atomic, Molecular and Optical Sciences Program; Chemical Sciences Geosciences and Biosciences Division, DOE OBES (M.J.B) and the SLAC LDRD program (M.J.B, H.L); the US DOE through Lawrence Livermore National Laboratory under the contract DE-AC52-07NA27344 and supported by the UCOP Lab Fee Program (award no. 118036) and the LLNL LDRD program (12-ERD-031); the Hamburg Ministry of Science and Research and Joachim Herz Stiftung as part of the Hamburg Initiative for Excellence in Research. We also want to thank the National Science Foundation for

providing funding for the publication of this work through the BioFEL Science Technology Center (award 1231306).

We thank H. Isobe, M. Shoji, S. Yamanaka, Y. Umena, K. Kawakami, N. Kamiya, J.R. Shen and K. Yamaguchi for permission to show a section Fig. 6 of their publication(Isobe et al., 2012) in Fig. 3D of this publication.

We thank Richard Neutze and his team for support and very helpful discussion during joint beamtime for the PSII project and his projects on time-resolved wide-angle scattering studies. We thank Axel T. Brunger for very helpful discussions concerning data analysis. We thank Thomas Terwilliger for support with parameter setting of phenix.autobuild program for the SA-omit maps. We also wish to thank R Burnap for helpful discussion concerning interpretation of results of ligand mutagenesis. We thank James D. Zook for his contributions concerning plastoquinone quantification. We thank Mary Zhu for helping to create high resolution figures for this publication. We thank Raytheon Ltd. for support of our studies by providing Night-vision devices.

Chapter 4

Time-Resolved Femtosecond Nanocrystallography of Photosystem II at Higher S-States

Introduction

This chapter describes time-resolved SFX experiments (TR-SFX) of Photosystem II with triple laser excitation, and therefore represents an extension of the experiments discussed in Chapter 3. The previously discussed S-state cycle involves four distinct charge-separation events, including the S_3 state, which is reached after two laser excitation flashes and first conformational changes have been detected. This chapter will discuss the experimental changes that took place in order to allow for visualization of the S_4 state with triple laser excitation and will examine preliminary data analysis results that have been obtained.

The first TR-SFX experiments were published in 2011 by Aquilla et al. (Aquila et al., 2012)using PSI-ferredoxin co-crystals as a model system. In this study, differences between the diffracted intensities of the excited and unexcited states were observed with time delays of 5μs and 10μs between laser excitation and interaction of the nanocrystals with the X-ray beam.

The kinetics of the S-state transitions are in the micro to millisecond time range ($S_0 \rightarrow S_1$ 30μs, $S_1 \rightarrow S_2$ 70μs, $S_2 \rightarrow S_3$ 190μs, $S_3 \rightarrow S_4$, 250μs $S_4 \rightarrow S_1$ 1100 μs) (Siegbahn, 2008; Yano et al., 2008) and changes in the Mn distances have been proposed based on detailed XAFS studies. Pump-probe time-resolved data on PSII have been collected

during three beamtimes at LCLS during August 2011, January 2012, and June 2012. At the August 2011 beamtime we achieved microsecond time-resolution by illumination of the crystal stream and synchronization of the optical laser and the FEL pulses, which explored of the $S \rightarrow S_2$ transition. During the beamtime in January 2012 we collected data of the $S_1 \rightarrow S_2 \rightarrow S_3$ transitions with two laser excitation pulses. Results of these experiments have been presented in Chapter 3 and published in Kupitz et al. In June 2012, we established a three laser excitation scheme, to advance the OEC in PSII to the transient S_4 state ($S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4$) (refer to Fig 4.1 for all timing schemes). This is the excited transient state where oxygen is evolved. The three laser excitation scheme requires two lasers. Laser 1 provides the light pulses one and two and is triggered by the incoming X-ray FEL pulse "1" while laser 2 provides the third light pulse, which is triggered by the previous X-ray-FEL "-1." This system allowed us to achieve the first data collection at the $S_4 \rightarrow S_0$ transition. While the system returns to the S_0 state in two ms, we have collected data 250 μs after the third optical laser excitation flash, to capture a time point of the transition S_4 state. Fig. 4.1 shows the laser excitation scheme for these experiments, which break new ground, as they are the first time that molecular structures can be determined in the transition from the transient S_4 state to the stable S_0 state. The final goal of future TR-SFX experiments is to explore the full dynamics of the S-state cycle and the oxygen evolution process.

In the original plan for the June 2012 beamtime we hoped to collect data at three different time points. However, due to experimental limitations with sample delivery and hardware at the beamline, we focused our data collection at the 250 μs time point.

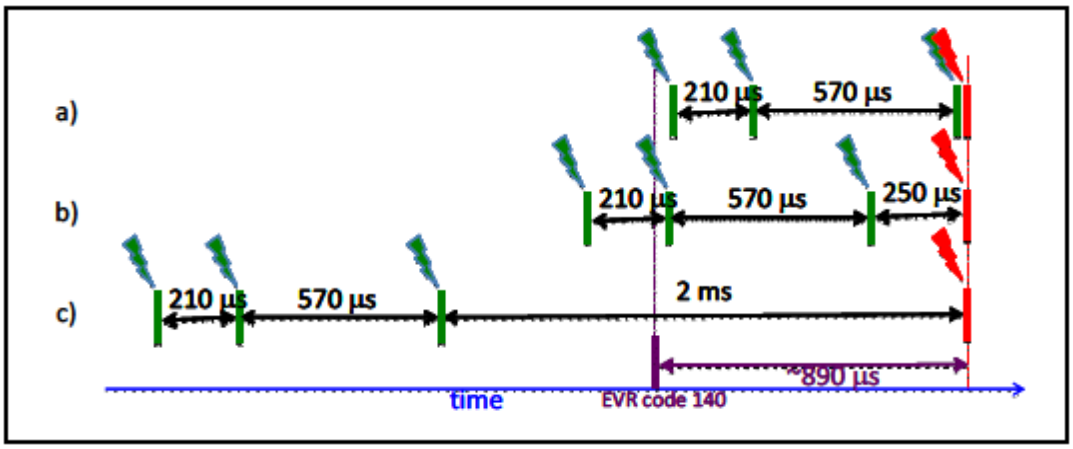


Figure 4.1. Laser excitation scheme of S_4 *that was used for the pump probe experiments on Photosystem II crystals. (a) time delays for fast time scales <1 µs,(b) time delays up to 250µs to reach the S_4 state, and (c) long time delays of up to 2ms would allow us to monitor the $S_4 \rightarrow S_0$ state transition.*

Methods

In order to reach the S_4 state, changes were required in the sample preparation and crystallization procedures. The most important of these adaptations was that a quinone was necessary to enable the third charge separation event to proceed. Fig. 3.1a shows the S-state transition scheme. In each charge separation excitation event, one electron is transferred from P680* to the mobile plastoquinone PQ_B in the Q_B binding pocket. This means that after two saturating laser flashes the double reduced PQ^{--} takes up two protons and leaves the Q_B binding pocket as plastoquinol PQH_2. In the native membrane PQH_2 is replaced by a PQ from the PQ pool located in the membrane.

Oxygen Evolution Rates in the Presence of Quinone Derivatives

For the system to proceed further than the S_3 state in the S-state cycle, an electron acceptor must be present. This is because after two excitation events, the quinone leaves and another must take its place. See Fig. 3.1a for a detailed view of the S-state cycle.

Therefore, we have tested several potential quinone derivatives for their ability to maintain the photocycle in the isolated PSII and its crystals. We measured the rate of oxygen evolution on re-dissolved crystals of Photosystem II to determine which quinones could maintain fast steady state oxygen evolution rates. The results are shown in Table I.

Table I. Oxygen Evolution Comparison

Name	Rate of Evolution (μmol Chl/mL/min)
PQ_9	0.14
Tocopherylquinone	0.0644
Benzoquinone	0.196
Vitamin K1	0.0255

The other quinone selection criterion was how closely the structure resembled that of the native plastoquinone. After initially comparing several commercially available quinones we were disappointed that none of the commercially produced quinones both resembled the native quinone, and yet had a fast O_2 turnover. With this realization, we initiated a new collaboration with Drs. Tom and Ana Moore with the goal to synthesize a synthetic quinone which would more closely resemble the native plastoquinone. Dr. Jesse Bergkamp derived the synthetic quinone (PQ_9). This quinone contains the same head-group as the native plastoquinone but features a shorter tail. Instead of the long isoprene tail of the native PQ, PQ_9 contains a simple nine carbon tail. See Fig. 4.2 for a comparison of the native vs synthetic quinones. As shown in Table I, the oxygen evolution rate determined is high for PQ_9 and the small benzoquinone.

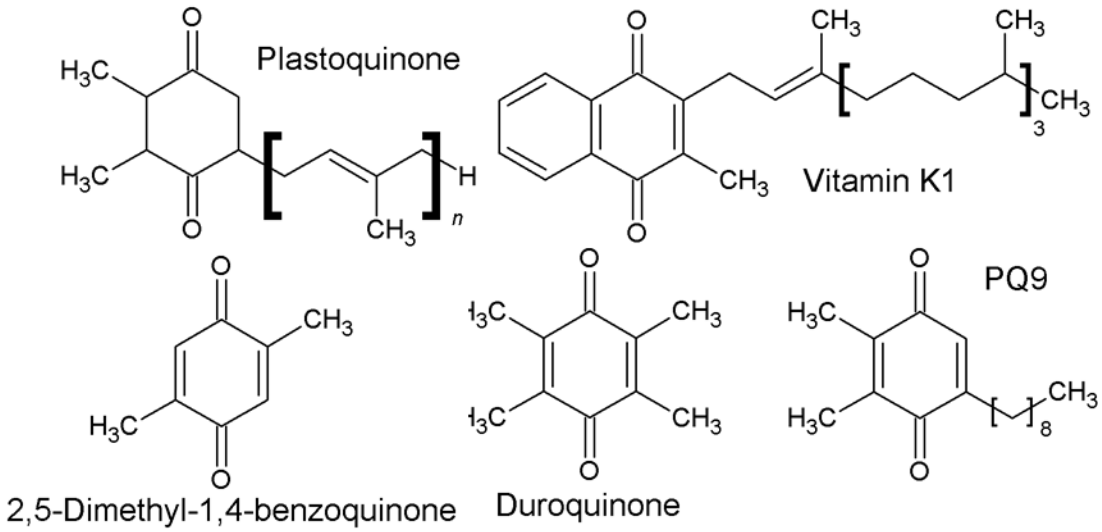


Figure 4.2. Quinone Structure for Electron Acceptor Replacement. *The structure of various quinones tested for their function as an electron acceptor for PSII during the crystallization process. For comparison, the structure of the native plastoquinone is also shown (top right).*

Crystallization in the Presence of Quinone Derivatives

As PQ_9 more closely resembles PQ than benzoquinone, crystallization experiments were performed in the presence of PQ_9.We found that we could apply the same FID methods for the sample with and without addition of PQ_9. See Chapter 2 for a review. The PQ_9 was dissolved in DMSO to make a stock solution of 0.5 M. This stock solution was added to crystallization buffer C leading to a final concentration of (5mM MES pH 6.4, 10 mM $CaCl_2$, 10 mM $MgCl_2$, 0.03% βDDM, and 0.2mM PQ_9). This solution intentionally contains an excess of plastoquinone to protein (33 PQ_9/1 P680) to ensure quantitative binding and full occupation of the Q_B binding pocket before PSII is excited by the third laser excitation flash. Thereby, the system mimics the quinone pool

that is found in the native membranes. Other than this minor change, no other changes were made between the preparation of the crystals used for the S_3 (double-flash) and S_4 (triple-flash) TR-SFX experiments.

Collection and Evaluation of Triple-Excited TR-SFX Data

The TR-SFX data collection on the triple-excited state used identical methods and experimental parameters as those described in Chapter 3, except for the change in the laser excitation scheme and the addition of the PQ_9 as described above. Data analysis is still ongoing, and therefore writing an in-depth procedure at this time would be premature. However, the current methods described in Chapter 3 are currently being applied to the S_4 data set. The data statistic is shown in Appendix A Table A.III.

Results and Discussion

PQ_9 inclusion into the crystals was achieved after optimization of conditions. Crystals grew at the standard conditions that have been established for the double and single flash experiments at other beamtimes, i.e. crystals were grown in the presence of 13% PEG2000 at 0.5mM Chl. These experiments collected 33,373 diffraction patterns from the dark state data set and 32,190 diffraction "hits" were identified from the triple-excited state data set, leading to 11,664 indexed patterns for the dark state and 12,543 indexed patterns for the triple-excited state. The most astonishing result was that the unit cell constant change observed with double-excitation was reversed after triple-excitation, leading to a similar unit cell constant for the dark and S_4 states. Furthermore, the diffraction quality of the PSII nano/microcrystals that were grown in the presence of PQ_9 was significantly better than the diffraction quality of the crystals grown without addition

of quinones. The resolution of the diffraction patterns extended to 4 Å, while data quality of the crystals grown without quinone addition was limited to 5 Å resolution. This diffraction improvement may be caused by a stabilizing effect that PQ_9 could have on the conformation of the acceptor site of PSII, which allows for achievement of better crystal order.

Structural results are currently limited, due to the aforementioned fact that the analysis is still in the preliminary stages. The total number of indexed diffraction patterns from the S_4 state is only 12,543, which is borderline for the determination of accurate structure factors with high multiplicity. Even though structures have recently been published by other groups with much lower multiplicity than the multiplicity of our data, (Kern et al.) we would like to complete the data set by reaching higher multiplicity before the current preliminary results are published. For current statistics see Appendix A Table A.III.

Preliminary Structural Model of the Triple-Excited Transient S_4 State

Figure 4.3 presents the first preliminary electron density map derived from the S_4 data set. The phases were determined and the structure was solved by molecular replacement as previously described for the S_3 state. (Kupitz et al.) The electron density map at 1.5 sigma is shown at a resolution of 4.0 Å, which represents an increase in resolution by 1 Å over the previous experiments. The electron density map is also very well defined, as first side chains are already visible. A further indication of the quality of the map is the fact that the density for the Pheophytin (chlorophyll that lacks the central Mg^{2+}), which acts as the secondary electron acceptor, shows a "hole" in the density indicative of the lack of a central metal. Most exciting is the fact that the electron density

map allows for the identification of the quinones in the Q_A and Q_B PQ binding sites. While the Q_A site is occupied by a phylloquinone molecule (indicated by its long isoprenoyl tail), the electron density in the QB binding pocket indicates the presence of PQ_9 in the Q_B binding pocket.

This finding is a clear indication that the native PQ has left the binding pocket after the second laser flash, thereby allowing PQ_9 to enter the Q_B binding site. This means that in a significant portion of the crystals the S_4 state may have been reached after triple-laser excitation.

Currently, we are exploring DEN refinement and a complete set of omit maps to further improve the electron density map and optimize the refinement statistics.

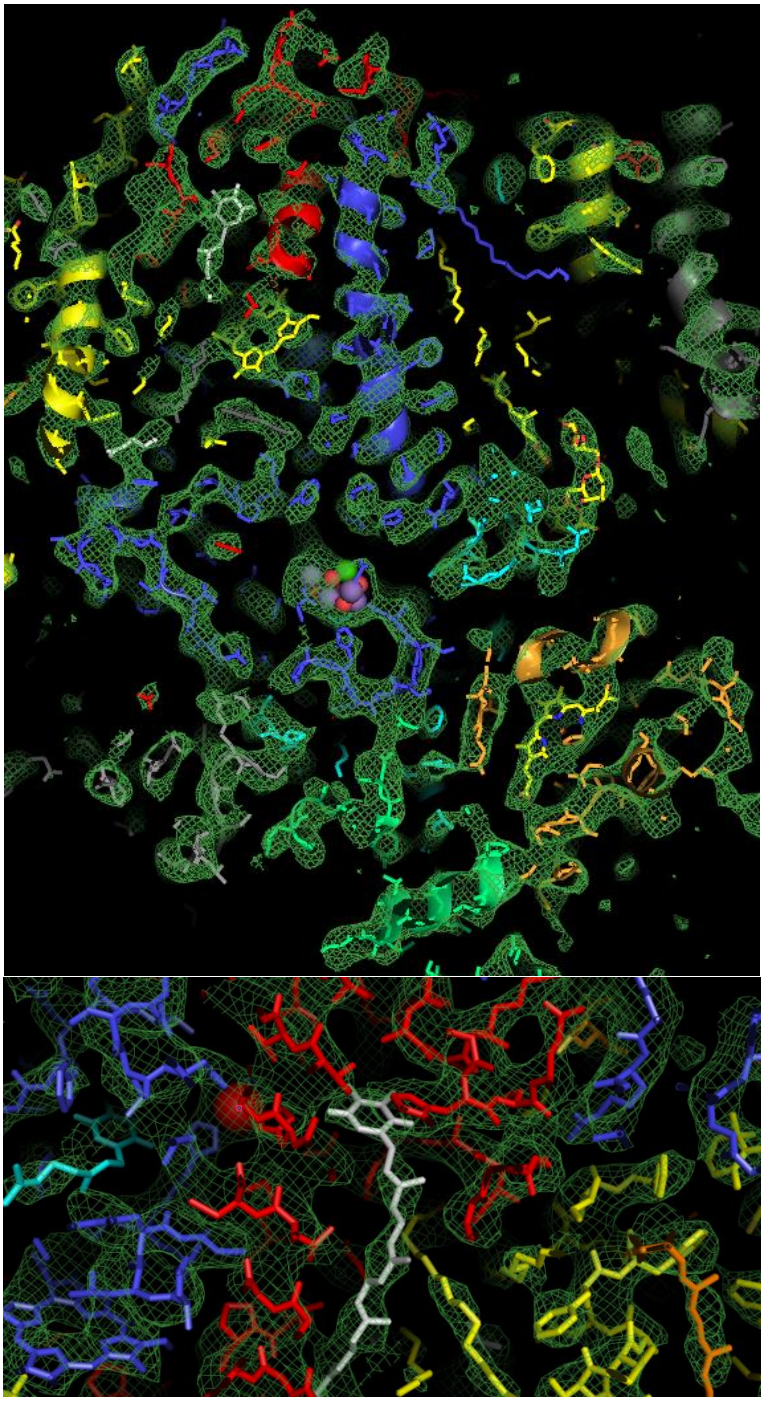


Figure 4.3. ***SFX electron density maps of the S_4 state*** *at 4.0 Å. (top) overview of PSII at 1.5σ showing good electron density correlating to the modelling map. (bottom) electron density map of the Q_a/Q_b binding region of PSII at 1 σ. Interesting to point out is that the density of the quinone in the PQ_B site on the left side of the image is truncated significantly shorter than a normal plastoquinone, this is an indication that it is the significantly smaller PQ_9 replacement.*

Chapter 5

Application of New Nanocrystallization Tools for membrane protein crystallization

Introduction

Previous chapters of this dissertation have described multiple crystallization techniques that typically require a precipitant solution that has a significantly higher density than the protein solution. If this is not the case, the interface and mixing zones will not form. This chapter describes the development of alternate nanocrystallization tools for the goal of establishing techniques that are broadly applicable to a large range of protein crystallization conditions. This is especially important for proteins that crystallize under conditions where there are no large differences in the density of protein and precipitant. One of the most prominent examples where FID cannot be applied is the salting-in crystallization, in which the salt must be removed in order for crystals to form(Fromme, 1996). Photosystem I is one example of a protein that crystallizes by decrease in ionic strength (Fromme, 1996). In order for these proteins to be successfully crystallized a new technique has to be developed, which ideally could be broadly applied to a whole range of different crystallization conditions, but could undergo many different variations in order to be optimized for a specific protein. Below, I will provide some background information that is important to understand the boundary conditions for these method developments.

Membrane proteins are extremely important for biology and medicine. As previously stated, over 60% of drug targets are membrane proteins, and approximately 30% of all human proteins are membrane proteins. They are difficult to express, isolate,

and crystallize as they are amphiphilic in nature and embedded into biological membranes. They have to be isolated and crystallized in form of detergent-protein micelles, which complicates the phase diagram and limits the precipitants that can be used for crystallization of the protein-detergent micelles in solution and in membrane mimics like the lipidic cubic phases (Cherezov et al., 2001; Caffrey, 2009; Caffrey and Cherezov, 2009). SFX allows for collection of high resolution X-ray diffraction patterns from nanocrystals, which are delivered to the X-ray beam in a fully hydrated stream at room temperature, while outrunning the radiation damage by limiting the X-ray pulse duration to femtoseconds. Membrane proteins have traditionally been notoriously difficult to crystallize. Less than 500 unique membrane protein structures have been determined to date while the total number of protein structures in the protein data base has exceeded 100,000.

Our team was able to be awarded beamtime at the LCLS in Feb 2012 to explore the use of SFX for nanocrystallization of membrane proteins. The screened proteins included targets from the ASU center for Membrane Protein of Infectious Disease Center (MPID), crystals of transporters from UCSF (group of R. Stroud), and crystals of a membrane protein from the group of W. Cramer at Purdue University. While it was originally planned that our collaborators would provide the nanocrystals of their proteins for the beamtime, which would be crystallized in the corresponding laboratories, results were inconsistent and we decided that we would try to crystallize all proteins on site at LCLS. This was a real challenge as we were asked to create nanocrystals from nine

proteins in under a week without having nanocrystallization conditions established previously from these proteins. Based on the information in the previous chapters, this would be considered an impossible task.

Methods

The proteins provided for these experiments were: CapA from *F. Tularensis*, IMPDH, C terminal tail of the corona virus M-protein, ABC transporter, purple bacteria RC-LHC1 complex, fusion protein CTB-MPR (HIV membrane protein gp41), OmpFT83, TolC-N190, and Rep 24. Many of these were samples that our lab had never worked with, and while rudimentary crystallization conditions had been identified, no nanocrystallization conditions had been established prior to the experiment.

As discussed previously, the time it would normally take to screen for conditions in which the proteins crystallized was simply unrealistic. Therefore, a new method had to be adapted that would allow for a high chance of success for crystallization, even if it could not be optimized. To do this, the supersaturated phase in the phase diagram was approached in a different method than the previous techniques, using a concentration technique that had originally been developed for crystallization of Photosystem I at low ionic strength (Hunter et al., 2011), but had not been tested with any protein crystallized under high salt or PEG conditions. Three typical precipitants were chosen: Sodium Malonate, Sodium Citrate, and Sodium Formate. These precipitants were all prepared at high concentrations, typically around 4 M in concentration in the buffer used for the isolation of the corresponding proteins. These solutions were adjusted according to the pH of the provided proteins, typically between a pH of 6-9.

Once the solutions were set, we attempted a concentration technique. To start the experiment, 150 µL of protein were mixed with 250 µL of one of the precipitant solutions. This was the placed into a microconcentrator with a molecular weight cutoff corresponding to the protein and detergent micelle size. Two different concentrators were used: slope-sided and flat-bottom concentrators were tested. The proteins were then concentrated at a very rapid pace, to achieve a visible pellet inside of the concentrator.

When this had been performed, the remaining volume (typically 50 – 80 µL) was used to wash the filter and re-suspend the pellet, in order to retain as many crystals as possible. A 2 µL drop was then taken from each experiment and tested for existence of nanocrystals using an optical microscope, SONICC, and DLS. Size, crystallinity, and general shape, as well as an estimate of crystals density was gathered from these measurements. These crystal suspensions were then freshly used to collect SFX diffraction patterns during the LCLS beamtime.

Results and Discussion

The results of the screening of these crystals will be summarized below. Most of the data will not be shown as the extensive data evaluation and comparison results are still in progress. The inclusion of the data in this thesis would be premature and require permission from many collaborators. The focus of the results will primarily be demonstrated with CTB-MPR because a publication has already been accepted. (Lee et al., 2014).

Each of the three precipitant solutions used worked to varying degrees, with the most positive hits achieved with Sodium Malonate as the precipitant. With this unified approach we were able to obtain nanocrystals of all nine proteins. It was important for the

success that precipitant solutions matched the protein buffer in salt, buffer, and detergent compositions, as the samples where these conditions were not met precipitated irreversibly. Citrate and Formate were both capable of crystallizing several of the proteins, but due to limitations in beamtime we could not test the diffraction quality of all samples. We therefore decided that the samples crystallized with Malonate would have the highest priority for data collection. This gave us the opportunity for a direct comparison between the crystals grown all with the same technique. Malonate was also chosen because at the time when crystal growth was first observed it was impossible to tell which crystals might diffract the best.

This technique was successful due to its approach to reach the supersaturated phase by increasing the protein concentration at constant precipitate concentration to reach the supersaturated zone of the phase diagram. The typical approach to reach the nucleation zone of phase diagram is to concentrate both protein and precipitant slowly by vapor diffusion. Yet, with this method, the rate of reaching the supersaturation zone depends on the volume; for small drops the volume to surface ratio is high, so equilibration is achieved quickly. The increase of the volume of the protein drop decreases the surface to volume ratio leading to slower rate of increase of supersaturation. The slower supersaturation leads to a decrease in nucleation rate while crystals have more time to grow slowly. While this is beneficial for growth of larger single crystals it hinders the formation of nanocrystals in larger batch VD experiments under conditions where nice nanocrystals have been observed before in small scale experiments. This makes any optimization of vapor diffusion experiments for SFX an elaborate task, where 1 ml of crystal suspensions has to be gathered from several hundreds of small crystallization

drops. By using this new technique and picking a reasonable concentration of precipitant, a fast "run" through the phase diagram is achieved. The protein concentration is constantly increased until the nucleation zone is reached, see Fig. 5.1. Faster spin cycles typically correlate to the production of a larger number of smaller crystals, while fewer but slightly larger crystals are observed at lower centrifugation speeds. Thereby, if the sample is spun slower, then it is possible to create fewer nuclei, which will lead to the growth of larger crystals. However, if the sample is spun too quickly, a mixture of both nanocrystals and amorphous precipitate is formed.

The main difficulty with the use of this technique for SFX is that it can be very difficult to create enough protein crystals and a sufficient crystal density to run an entire beamtime, as the final volume is limited by the concentrator. It requires milliliters to run an entire beamtime, compared to only microliters per experiment yielded with this technique. There is the possibility to scale this technique up using larger concentrators, however at this time the experiments hadn’t been performed.

In the experiments, two separate types of microconcentrators were tested: those with the filter flat on the bottom, which makes it possible to run the filter entirely dry, and concentrators with the filter along the sides of the concentrator and a small reservoir at the bottom, which makes it impossible to completely drain the concentrator. There are advantages to both set-ups, however strictly from a crystallization point of view the flat-bottom filters tend to achieve the most promising results. We found that the superior results with the flat bottom filter are achieved because the crystals tend to nucleate on the filter, and with the filter on the bottom all of the crystals remain at the filter surface which is fully submerged in solution, leading to a further increase in nucleation. However, with

the filters on the sides the majority of crystals forming on the filters stick to the filters and they remain even after the sample volume has decreased, thereby they "stick" to the filter on the sides of the concentrators and subsequently "dry out." Hence, a considerable amount of crystallites is lost leading to lower crystal density. However the flat-bottom concentrators must be constantly monitored so that not all solution runs through the filter, which lead to drying of the crystals thereby ruining the entire crystallization batch.

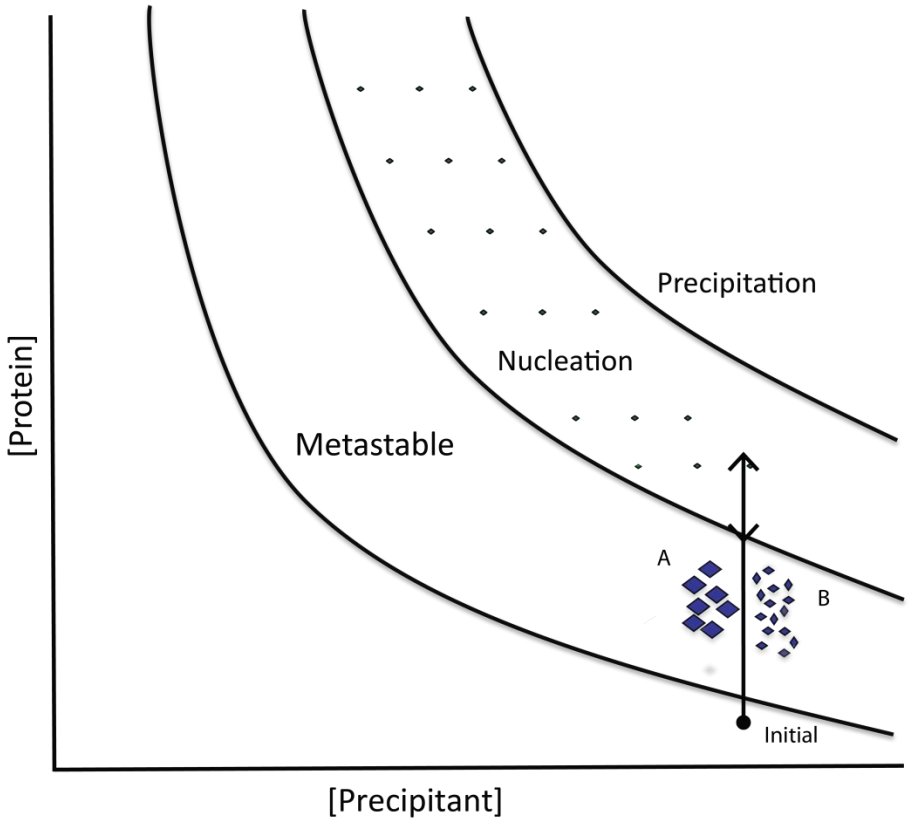


Figure 5.1. ***An example of a general phase diagram*** *with the centrifugation concentration techniques approach to the phase diagram drawn in. Starting at a high precipitant and low protein concentrate the protein into the nucleation phase, as crystals form the protein is depleted but concentration continues to increase yield thus the double arrow. (A) Larger crystals which are grown by slower spinning thus creating fewer nucleation points and allowing the crystals a longer chance to grow. (B) As seen from the arrow the longer the protein is in the nucleation zone, and the further in it is the more nucleation you get thus the larger amount of smaller crystals you can obtain.*

Of the crystals made, for eight out of nine single crystals X-ray diffraction patterns were observed. All of the samples listed above had indexable diffraction patterns. For four of these proteins the SFX X-ray diffraction data were the first X-ray diffraction patterns ever recorded. Fig. 5.2 and Fig. 5.3 show examples of crystals and diffraction patterns collected from CTB-MPR, just one of the proteins which was tested successfully at this beamtime. The paper with these results is in press (Lee et al., 2014). It should be noted that crystallization conditions were available for this protein, so the sample was crystallized in 30% PEG3350, 0.2M NH_4HCO_2, 0.01M $CaCl_2$, 0.05 M HEPES pH7.5, and 0.02%βDDM. Data statistics for the CTB-MPR data collected at LCLS can be viewed in Table II.

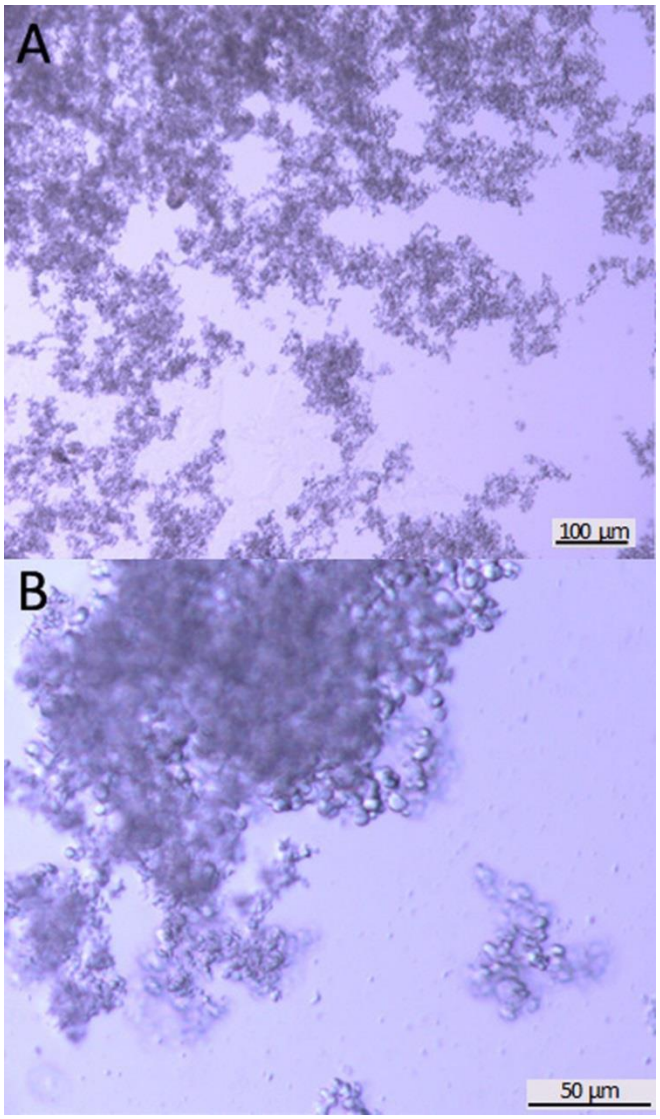


Figure 5.2. ***Nano/microcrystals of CTB^{AAAA}MPR*** *grown in 0.2 M ammonium formate and 30% PEG3350 before (A) and after (B) filtering through a 20 µm filter, shown at a higher magnification. (taken from (Lee et al., 2014))*

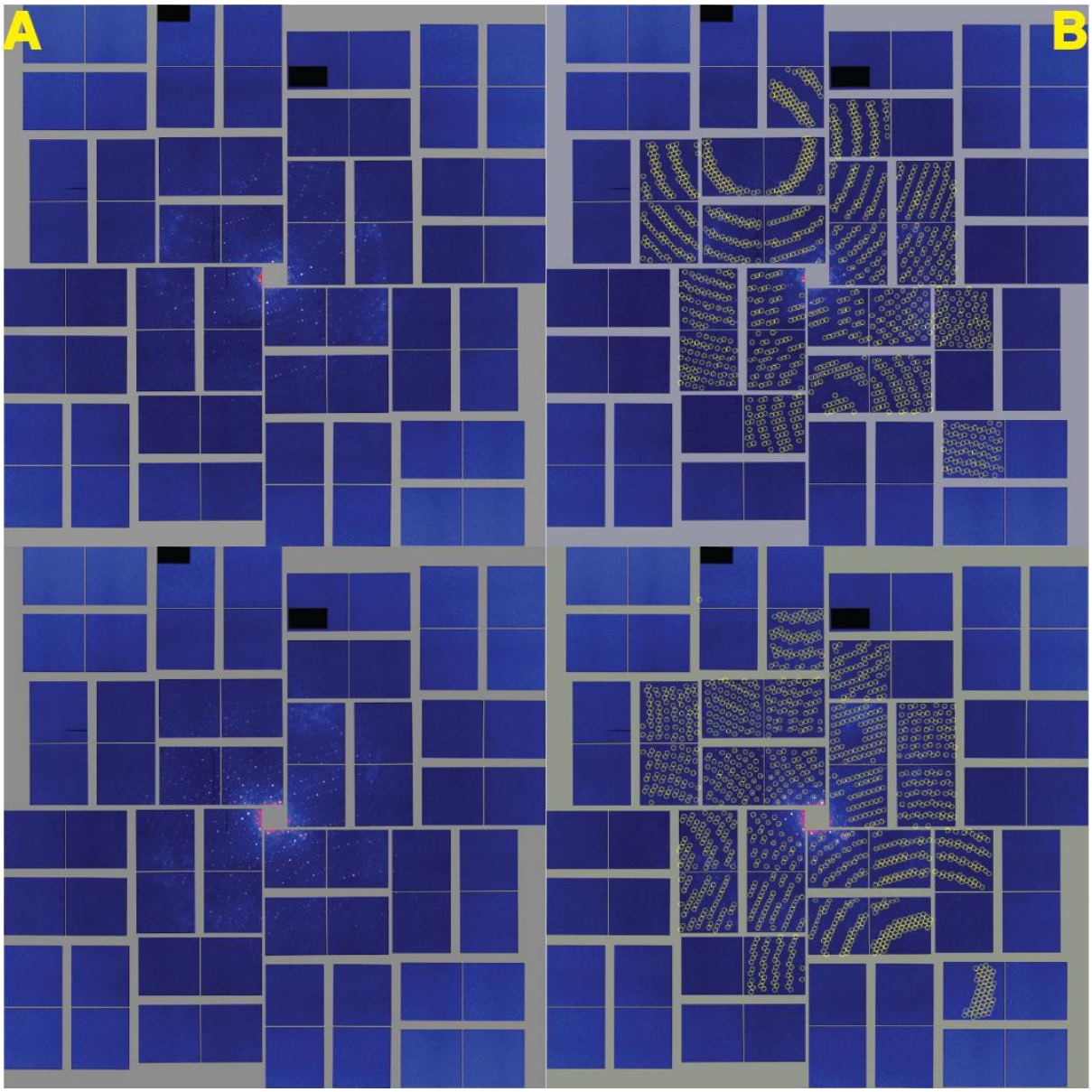


Figure 5.3. ***CTB-MPR Diffraction Patterns*** *(A) Two CTB-MPR diffraction patterns collected from nano/microcrystals at the CXI beamline at LCLS in Feb 2012. (B) Indexing of the diffraction patterns in (A) (taken from (Lee et al., 2014))*

Table II: Crystallographic Data for CTB^{AAAA}MPR at LCLS. Taken from (Lee et al., 2014)

Run Time	10 min 40 sec
Total number of raw frames	72, 767
Number of crystal hits	1,006
Hit rate	1.38%
Number of indexed patterns	55
Indexing yield	5.46%
Unit cell	a=b=c=332 Å; α = β = γ = 60°
Space group	R32

Since the concentration technique was first attempted, it has been tested on many other proteins with various degrees of success. Examples of successful crystallization include the Eukaryotic Complex I and the soluble protein Photoactive Yellow Protein (PYP). Optimization of the technique with establishment of the phase diagram has drastically increased the crystallization yield, thus making it a viable technique for providing large amounts of sample that could sustain an entire beamtime.

Chapter 6

Outlook and Conclusions

Even though X-ray protein crystallography is the primary method of solving protein structures, there are still many difficulties associated with the crystallography, as discussed in the previous chapters. There are two main difficulties of traditional macrocrystallography: the need for large well-ordered protein crystals, which can take years to optimize, if growing conditions can even be found, and the reduction in quality which occurs due to X-ray radiation induced damage. Both of these problems are solved by the technique of femtosecond nanocrystallography.

The time-resolved Photosystem II project was started in an attempt to develop methods that could unravel the conformational changes that occur in PSII while it performs light absorption and electron transfer coupled to the oxidation of water, leading to oxygen evolution. The ultimate goal of the project is to be able to make a molecular movie of the conformational changes by taking the snapshots of the protein at different S-states and different time points in the transition between these S-states. Multiple initial goals have to be reached in order accomplish this project, including improving the protein yield of the traditionally meager PSII protein preparations, determining the best crystallization conditions and techniques for the growth of nano and microcrystals, and finding proper quinone type electron acceptors that can replace the natural plastoquinone after it leaves the binding site as plastoquinol in order to advance PSII through all the S-states. Additional experiments were needed to for biophysical characterization of the

PSII samples. Furthermore, it was very important to find new ways to improve the resolution of the crystals, and improve the reproducibility of all steps in the process from cell growth through isolation and crystallization of PSII in order for this project to be a success. In addition, many other experiments not discussed in this dissertation were required, such as contributions to laser timing development and data analysis work.

Results Summary

In this project, massive amounts of data have been generated, both during the course of crystallization experiments, but particularly during the course of the time-resolved SFX experiments at the Free Electron Laser LCLS.

This project has shown the first actual conformational change recorded in the S-state cycle, which is also the first study that shows the proof of principle for TR-SFX. It thereby paves the way for studies on other photoactive proteins and may lead to the determination of molecular movies in the future. This success was possible due to the many groundbreaking new features developed for the advancement of the SFX technique. The detection of conformational changes was achieved by taking diffraction snapshots of PSII crystallites in the dark and after laser excitation, and finally led to the determination of electron density maps for the laser excited and dark states. We also collected the first damage free time-resolved data set of PSII, which observes conformational changes in the highly damage prone OEC.

The crystallization experiments described in this dissertation are highly relevant to the field of SFX, as there have not been extensive studies of crystallization techniques for micro/nanocrystals in the past. These techniques have been successfully applied to multiple proteins at beamtimes at LCLS, providing evidence that they can be efficiently

and effectively used for the microcrystallization of a large range of proteins for SFX use, which opens the field further for the structure determination of complex proteins from which very few structures have so far been determined.

Summary of Experiments

The development of SFX involves a large international collaboration, where our group works with teams across the globe hand in hand on new discoveries. The expertise that I have gained during the work on TR-SFX studies on Photosystem II spans the wide range from protein isolation, technique developments for nanocrystal growth and characterization, to sample delivery and collection of X-ray diffraction data at LCLS. With this field being so young, very few groups worldwide have expertise in SFX and therefore we established collaborations with many groups on the application of SFX methods to their important biological problems. I have therefore collaborated during my dissertation with more than a hundred scientists worldwide, which included their visits to our lab where I performed crystallization experiments with them and worked with them on crystal characterization including SONICC and DLS experiments. The collaborative work also included joint experiments at LCLS, where crystals were grown and characterized on site, with SFX data collected and evaluated in a joint effort, leading to ten plus joint publications. To include a full comprehensive description of each of the projects would have by far extended the scope of this thesis. I have therefore in this section included the abstracts from select resulting publications, as well as a brief summary of the highlights of the results of some of most significant collaborative experiments and results that have been accomplished during my doctoral studies.

High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography

Boutet S. et al. (2012) High-resolution protein structure determination by serial femtosecond crystallography. Science **337:** 362-364

Abstract: Structure determination of proteins and other macromolecules has historically required the growth of high-quality crystals sufficiently large to diffract x-rays efficiently while withstanding radiation damage. We applied serial femtosecond crystallography (SFX) using an x-ray free-electron laser (XFEL) to obtain high resolution structural information from microcrystals (<1x1x3 μm^3) of the well-characterized model protein lysozyme. The agreement with synchrotron data demonstrates the immediate relevance of SFX for analyzing the structure of the large group of difficult-to-crystallize molecules.(Boutet et al., 2012)

This paper reported the first proof-of-principle demonstrating the SFX can achieve X-ray diffraction from proteins extending to high resolution using lysozyme as a model system. The work also included a damage study showing that the structure determined from SFX data represents an almost exactly similar structure compared to a low radiation dose lysozyme structure that was obtained from large crystals using synchrotron data, where data were collected on crystals (prepared with similar crystallization conditions). Both data collection at LCLS and at the Synchrotron were performed at room temperature. The set-up of the experiment can be seen below in Fig. 6.1. The radiation dose limit typically set for macrocrystallography is considered to be 33 MGy. The dose absorbed by the large crystal during data collection was only 24 kGy, well under the normal limit. As Fig. 6.2a shows, an electron density map can be calculated at high resolution. The SFX data collected with different pulse duration times from 5-40 fs were compared to synchrotron data, (see Fig. 6.2b). Interestingly, no interpretable differences in the electron difference map were observed, even when SFX data were collected using longer pulse lengths for SFX, showing that the diffract before destroy principle extends to high resolution.

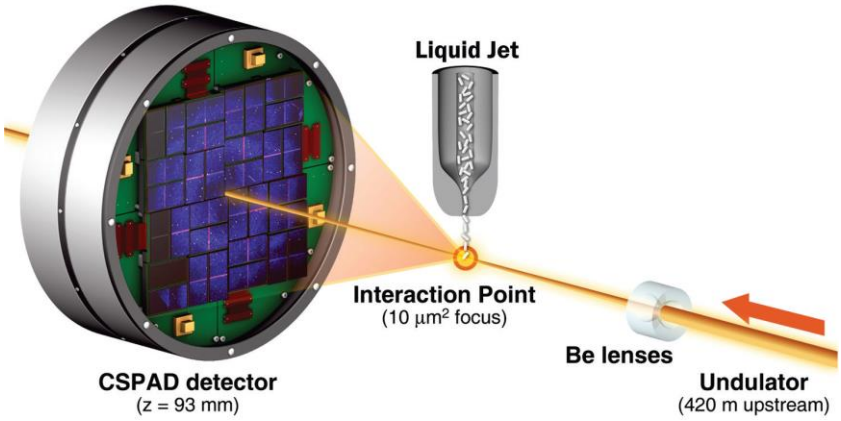


Figure 6.1. ***Experimental geometry for serial femtosecond crystallography*** *at the Coherent X-Ray Imaging instrument. Single pulse diffraction patterns from single crystals flowing in a liquid jet are recorded on a CSPAD at the 120 Hz repetition rate of LCLS. Each pulse was focused at the interaction point using 9.4 keV x-rays. The sample-to-detector distance (z) was 93 mm.(Boutet et al., 2012)*

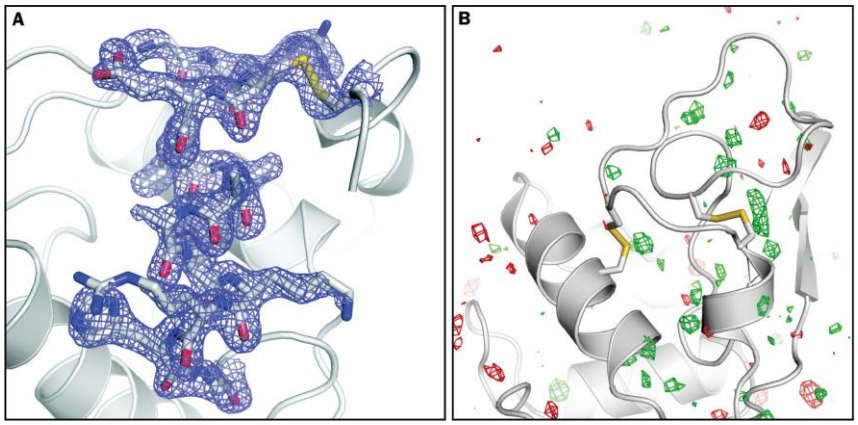


Figure 6.2. ***Final, refined*** *$2mF_{obs}$-DF_{calc} (1.5σ)* ***electron density map of lysozyme at 1.9 Å resolution*** *calculated from 40 fs pulse data. (B) F_{obs}[40 fs]-F_{obs}[synchrotron] difference Fourier map, contoured at +3 σ (green) and -3 σ (red). No interpretable features are apparent. The synchrotron dataset was collected with a radiation dose of 24 kGy.(Boutet et al., 2012)*

Natively Inhibited Trypanosoma brucei Cathepsin B Structure Determined by Using an X-ray Laser

Redecke L. et al. (2013) Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser. Science **339:** 227-230

Abstract: The Trypanosoma brucei cysteine protease cathepsin B (TbCatB), which is involved in host protein degradation, is a promising target to develop new treatments against sleeping sickness, a fatal disease caused by this protozoan parasite. The structure of the mature, active form of TbCatB has so far not provided sufficient information for the design of a safe and specific drug against T. brucei. By combining two recent innovations, in vivo crystallization and serial femtosecond crystallography, we obtained the room-temperature 2.1 angstrom resolution structure of the fully glycosylated precursor complex of TbCatB. The structure reveals the mechanism of native TbCatB inhibition and demonstrates that new biomolecular information can be obtained by the "diffraction-before-destruction" approach of x-ray free-electron lasers from hundreds of thousands of individual microcrystals.(Redecke et al., 2013)

These experiments provided detailed insight into the molecular basis of a very serious disease, sleeping sickness. They combined two novel techniques, including *in vivo* crystallization. In this method, instead of standard crystal growth, an insect cell expression system is used in which the protein is crystallized in the cell and then harvested directly from the cell for structure determination. These experiments were the first time diffraction from *in vivo* grown crystals was tested and observed with an XFEL and during the first attempt spectacular results were already obtained. An electron micrograph of the crystals inside the insect cell, as well as a cartoon diagram of TbCatB complex, can be seen in Fig. 6.3. From the data gathered here, a novel structure was determined of the CatB protease for the first time crystallized with the N-terminal propeptide included, that lies in the catalytic cleft. As shown in Fig. 6.4, new electron density for this propeptide was discovered, a feature that hadn't been included in any of the macrocrystallography structures of catB (Redecke et al., 2013). This propeptide

serves as a natural inhibitor, blocking the enzyme function when the protease is inside the parasite, but it is cleaved upon entry into the red blood cells, thereby dissolving them. This structure may thereby form the basis for a promising drug design against sleeping sickness. The map also confirms that certain regions, mainly loop regions, are involved in the binding process of CatB.

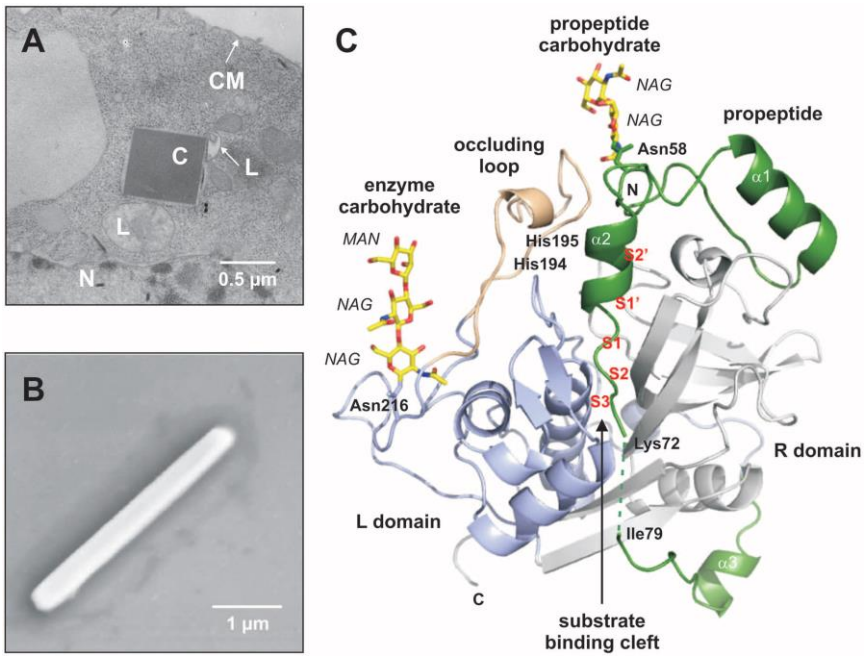


Figure 6.3. ***In vivo Grown Crystals and Three-Dimensional Structure of the TbCatB-propeptide Complex.*** *(A)Transmission electron microscopy (EM) of an infected Sf9 insect cell showing a crystal of overexpressed TbCatB inside the rough endoplasmic reticulum that is cut perpendicular to its long axis. N, nucleus; L, lysosome; C, crystal; CM, cell membrane. (B) Scanning EM of a single TbCatB crystal after isolation. (C) Cartoon plot of the TbCatB propeptide complex exhibiting the typical papainlike fold of cathepsin B–like proteases (supplementary text S1). Gray, R domain; blue, L domain; beige, occluding loop. The native propeptide (green) blocks the active site. The subsites of the substrate-binding cleft N-terminal (nonprime: S2, S3) and C-terminal (prime: S1′, S2′) to the active site (S1) have been identified by comparison with the human CatB structure (13) and labeled (red) according to Schechter and Berger (27). Two N-linked carbohydrate structures (yellow) consist of N-acetylglucosamine (NAG) and mannose (MAN) residues (yellow, carbon atoms; blue, nitrogen atoms; red, oxygen atoms).(Redecke et al., 2013)*

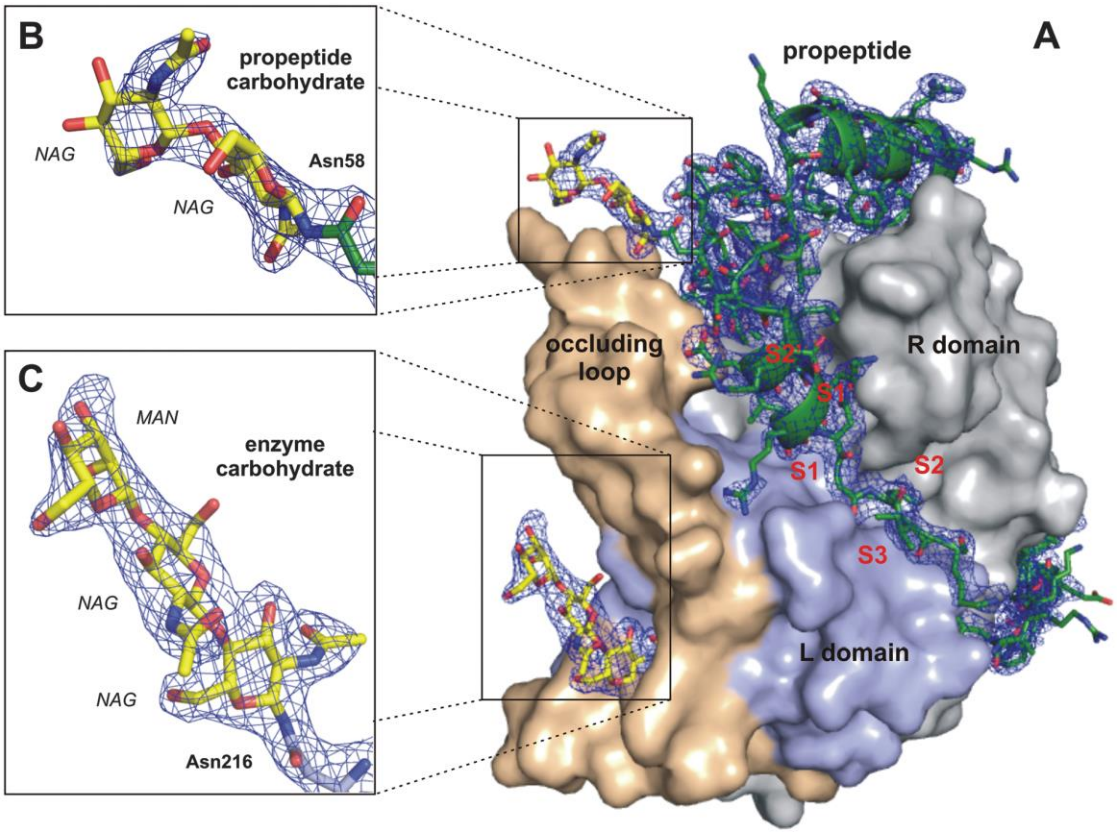


Figure 6.4. ***Quality of the Calculated Electron Density for TbCatB-propeptide.*** *(A) Surface representation of the TbCatB-propeptide complex solved by molecular replacement using the mature TbCatB structure (11) as a searchmodel. The solution revealed additional electron density (2Fobs – Fcalc, 1s, blue) of the propeptide (green) that is bound to the V-shaped substrate-binding cleft and of two carbohydrate structures (yellow) N-linked to the propeptide (B) and to the mature enzyme (C). The propeptide, as well as both carbohydrates, are well defined within the electron density map (blue), which confirms that the phases are not biased by the search model. Color codes correspond to Fig. 6.3(C).(Redecke et al., 2013)*

Serial Femtosecond Crystallography of G Protein-Coupled Receptors
Liu W. et al. (2013) Serial femtosecond crystallography of G protein-coupled receptors. Science **342:** 1521-1524

Abstract: X-ray crystallography of G protein–coupled receptors and other membrane proteins is hampered by difficulties associated with growing sufficiently large crystals that withstand radiation damage and yield high-resolution data at synchrotron sources. We used an x-ray free-electron laser (XFEL) with individual 50-femtosecond-duration x-ray pulses to minimize radiation damage and obtained a high-resolution room-temperature structure of a human serotonin receptor using sub-10-micrometer microcrystals grown in a membrane mimetic matrix known as lipidic cubic phase. Compared with the structure solved by using traditional microcrystallography from cryo-cooled crystals of about two orders of magnitude larger volume, the room-temperature XFEL structure displays a distinct distribution of thermal motions and conformations of residues that likely more accurately represent the receptor structure and dynamics in a cellular environment.(Liu et al., 2013)

This was the first time the lipidic cubic phase (LCP) was successfully used for data collection at an XFEL. For more details on LCP please see Fig. 6.5, as well as the reviews referenced here (Caffrey, 2009; Caffrey and Cherezov, 2009). The work is groundbreaking due to the fact that many membranes proteins that are very difficult to crystallize using regular crystallography in form of detergent micelles crystallize more readily in LCP, though growing them to sufficient size can still be problematic. They crystallize more readily because LCP is a native membrane environment which makes an ideal place for membrane protein crystallization. GPCRs are a class of protein that mediates cellular communication. Due to their prominent role in physiological processes (Liu et al., 2013), they are a very interesting target for drug design (Dong et al., 2013). However, for the aforementioned reasons, such as crystal growth, low resolution, and radiation damage, it can be very difficult to obtain structures of these proteins. This paper proved that by using small crystals in LCP high resolution membrane protein structures

can be determined. Moreover, this structure determined with SFX may be even more accurate than the original structure (Liu et al., 2013) obtained through macromolecular crystallography. This is due to the unique data collection method of SFX in which there is no damage and no freezing, which could bias the structure in a non-native way. The comparison that was done between SFX data and synchrotron data showed that several areas were differences between the two structures were observed which were attributed to the detrimental effects of freezing which are absent in the SFX data (See Fig 6.6). Even

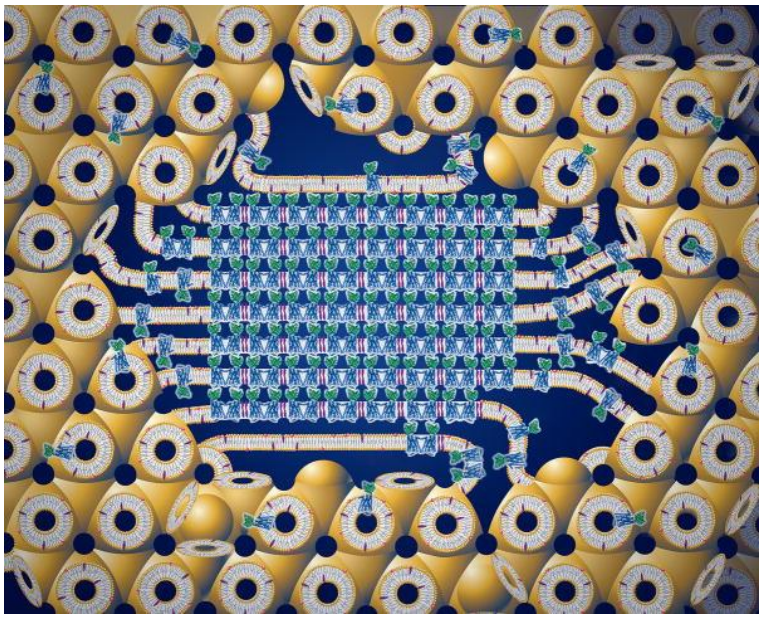


Figure 6.5. ***Cartoon Representation of LCP Crystallization****. Proteins are initially reconstituted into the curved bilayers of the cubic phase (tan). Added precipitants then shift the equilibrium away from stability in the cubic membrane. A phase separation occurs when this happens with protein molecules diffusing from the continuous bilayered reservoir of the cubic phase by way of a sheet-like or lamellar portal to lock into the lattice of the advancing crystal face (mid-section of figure). Co-crystallization of the protein with a native lipid (cholesterol) is shown here. The lipid bilayer is approximately 40 Å thick and as much as possible all molecules have been drawn to size. Image and caption taken with changes from (Caffrey and Cherezov, 2009).*

minute differences in the structure can be crucial for the understanding of the protein function and the dynamics of protein conformations can be critical for the understanding of the function and drug binding properties.

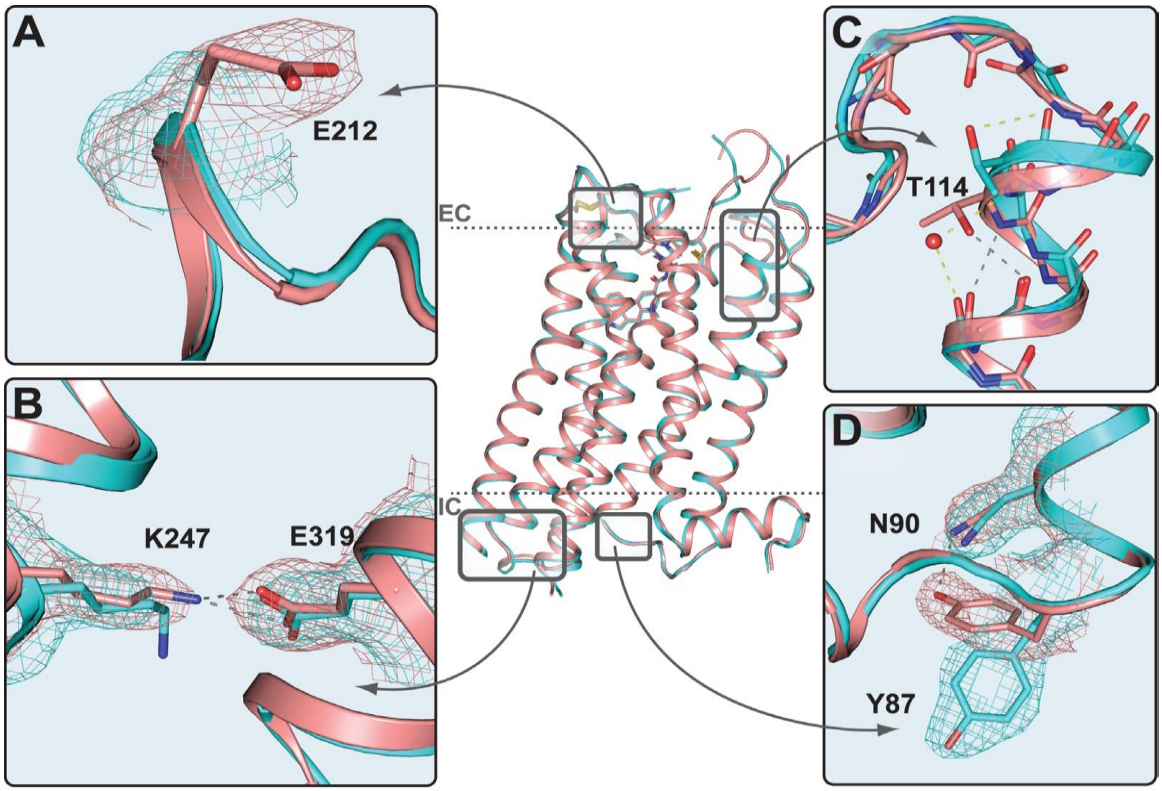


Figure 6.6. ***Comparison between 5-HT2B-XFEL (light red) and 5-HT2B-SYN (teal)*** *structures. Central image represents a backbone overlay of the two structures. Dashed lines correspond to membrane boundaries defined by the Orientation of Proteins in Membrane database (http://opm.pharumich.edu) (28). (A) Electron density for the Glu212 side chain is missing in 5-HT2B-SYN and fully resolved in 5-HT2B-XFEL. (B) A salt bridge between Glu319 and Lys247 links intracellular parts of helices V and VI in the 5-HT2B-XFEL structure. In the 5-HT2B-SYN structure, Lys247 makes a hydrogen bond with Tyr1105 from the BRIL fusion protein. (C) Extracellular tip of helix II forms a regular helix in 5-HT2B-XFEL with Thr114, making a stabilizing hydrogen bond with the backbone carbonyl, whereas in 5-HT2B-SYN, a water-stabilized kink is introduced at this position. (D) Tyr87 forms a hydrogen bond with Asn90 in 5-HT2B-XFEL; this hydrogen bond is broken, and Tyr87 adopts a different rotamer conformation in the 5-HT2B-SYN structure. 2mFobs-DFcalc maps (contoured at 1s level) are shown only around described residues.(Liu et al., 2013)*

Lipidic Cubic Phase Injector Facilitates Membrane Protein Serial Femtosecond Crystallography

Weierstall U. et al. (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. Nat Commun **5:** 3309

Abstract: Lipidic cubic phase (LCP) crystallization has proven successful for high-resolution structure determination of challenging membrane proteins. Here we present a technique for extruding gel-like LCP with embedded membrane protein microcrystals, providing a continuously renewed source of material for serial femtosecond crystallography. Data collected from sub-10 mm sized crystals produced with less than 0.5 mg of purified protein yield structural insights regarding cyclopamine binding to the smoothened receptor.(Weierstall et al., 2014)

This paper complements the research discussed in the previous section. This publication describes the operation principle and first application of the injector for delivery of crystals in LCP to the FEL beam. The injector is shown in Fig 6.7. The work also describes a structure on the Smoothened Receptor in complex with the drug cyclopamine, but in my opinion the most significant impact of this publication for the future of SFX is the immense benefit of using this type of injector. The largest problem that SFX faces is its extremely high sample consumption. By using this style of viscous jet, the sample consumption can be reduced from mLs to µL of sample per structure. See Table III for a comparison on sample consumption. In the case of the Smoothened Receptor, only 50 µL of sample were used to solve a structure of this protein. This is a drastic decrease in sample consumption which will make SFX available to many new applications that are currently sample-limited. It has also spawned a new sub-field, as work is in progress aiming to determine if a highly viscous LCP mimtic could be used to deliver any pre-grown crystal to the X-ray beam via a novel type of viscous jet.

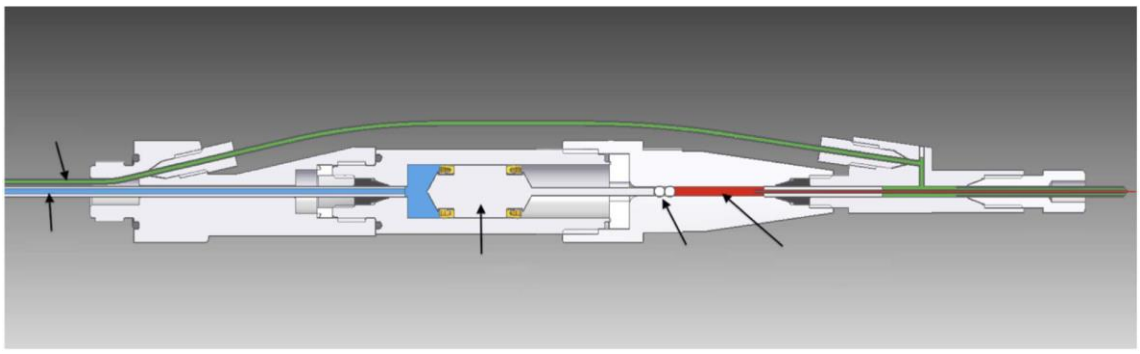


Figure 6.7. ***Schematic of the LCP injector****, middle section. In operation, the device is attached via the leftmost threaded fitting to a nozzle rod (not shown) for insertion into the experimental chamber. Water (blue) and gas (green) lines are routed through the nozzle rod from the left. LCP (red) is extruded from the nozzle on the right. Water, at a pressure of up to 300 psi, drives the hydraulic plunger, which amplifies the pressure 34 times to drive LCP through a capillary with an inner diameter of 10–50 mm. Two spherical Teflon beads are used to provide a tight seal against a pressure of up to 10,000 psi. The co-flowing gas is necessary for reliable extrusion and to maintain co-axial flow.(Weierstall et al., 2014)*

Table III. Sample Consumption and Hit Rates for Different SFX Experiments Comparing GDVN and LCP Injectors.

Protein	PS I [16]	Lysozyme [18]	5-HT_{2B}	DgkA
Crystal growth medium	Liquid	Liquid	LCP	LCP
Injector	GDVN	GDVN	LCP	LCP
Total frames recorded	1,850,000	1,471,615	4,217,508	1,987,632
Crystal hits*	112,725	66,442	152,651	263,435
Indexed hits	15,445	12,247	32,819	66,165
Hit rate, %	6	4.5	3.6	13
Indexing rate %	13.7	18.4	21.5	25
Sample consumption (liquid volume and protein mass)	10,000 µL 10 mg	10,800 µL 15 mg	100 µL 0.30 mg	46 µL 0.30 mg
Flow rate in □L/min	10	30	0.17	0.17
Consumed protein per 10,000 indexed patterns	6.5 mg	12.2 mg	0.091 mg	0.045 mg

For the PS I and Lysozyme datasets, hits where considered diffraction patterns with >10 Bragg spots, for the 5-HT_{2B} and DgKA datasets, hits are diffraction patterns with >15 Bragg spots. GDVN stands for gas dynamic virtual nozzle which is the typical liquid stream jet that SFX uses.(Weierstall et al., 2014)

Outlook

The first time-resolved structures of PSII presented here are the first proof of principle for the new method of TR-SFX and demonstrate the power of the time-resolved SFX technique. While the interpretation of the results is still preliminary, it produces

results which show great promise for further studies. The TR_SFX method can be extended to other photoactive proteins (for example photoreceptors) and with the use of caged substrates (for example caged ATP) even enzymatic reactions could be triggered by light. It allows also for the first time the study of irreversible reactions by crystallography. These advances are exciting news for the technique because it shows hope for time-resolved studies of many different proteins at work.

The true potential of SFX is shown in its ability to explore the spatial-temporal dimension of proteins. With new advances in the field, such as the lipidic cubic phase injection allowing the high-throughput study of medically relevant membrane proteins, as well the development of a mixing jet to allow for fast enzymological studies, the field of time-resolved SFX is in a unique position to expand our knowledge of protein biology at an extremely fast pace.

Many uses of SFX are still in their infancy, and it may yet be many years before time-resolved SFX has become a fully vetted technique. However, its future and the future of structural biology are very bright as more diverse and refined uses are continually being discovered. This is especially true since new free electron lasers are coming online in the next five years which will further increase their availability and continue to expand the user base, which will allow more highly interesting structural biology studies to be performed.

REFERENCES

Abdallah BG, Kupitz C, Fromme P, Ros A (2013) Crystallization of the large membrane protein complex photosystem I in a microfluidic channel. ACS Nano 7: 10534-10543

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallographica Section D-Biological Crystallography 68: 352-367

Altarelli M (2010) From 3rd to 4th-Generation Light Sources: Free-Electron Lasers in the X-ray Range. Crystallography Reports 55: 1145-1151

Altarelli M, Mancuso AP (2014) Structural biology at the European X-ray free-electron laser facility. Philos Trans R Soc Lond B Biol Sci 369

Ames W, Pantazis DA, Krewald V, Cox N, Messinger J, Lubitz W, Neese F (2011) Theoretical evaluation of structural models of the S2 state in the oxygen evolving complex of Photosystem II: protonation states and magnetic interactions. J Am Chem Soc 133: 19743-19757

Aquila A, Hunter MS, Doak RB, Kirian RA, Fromme P, White TA, Andreasson J, Arnlund D, Bajt S, Barends TR, Barthelmess M, Bogan MJ, Bostedt C, Bottin H, Bozek JD, Caleman C, Coppola N, Davidsson J, DePonte DP, Elser V, Epp SW, Erk B, Fleckenstein H, Foucar L, Frank M, Fromme R, Graafsma H, Grotjohann I, Gumprecht L, Hajdu J, Hampton CY, Hartmann A, Hartmann R, Hau-Riege S, Hauser G, Hirsemann H, Holl P, Holton JM, Homke A, Johansson L, Kimmel N, Kassemeyer S, Krasniqi F, Kuhnel KU, Liang M, Lomb L, Malmerberg E, Marchesini S, Martin AV, Maia FR, Messerschmidt M, Nass K, Reich C, Neutze R, Rolles D, Rudek B, Rudenko A, Schlichting I, Schmidt C, Schmidt KE, Schulz J, Seibert MM, Shoeman RL, Sierra R, Soltau H, Starodub D, Stellato F, Stern S, Struder L, Timneanu N, Ullrich J, Wang X, Williams GJ, Weidenspointner G, Weierstall U, Wunderer C, Barty A, Spence JC, Chapman HN (2012) Time-resolved protein nanocrystallography using an X-ray free-electron laser. Opt Express 20: 2706-2716

Aquila A, Hunter MS, Doak RB, Kirian RA, Fromme P, White TA, Andreasson J, Arnlund D, Bajt S, Barends TRM, Barthelmess M, Bogan MJ, Bostedt C, Bottin H, Bozek JD, Caleman C, Coppola N, Davidsson J, DePonte DP, Elser V, Epp SW, Erk B, Fleckenstein H, Foucar L, Frank M, Fromme R, Graafsma H, Grotjohann I, Gumprecht L, Hajdu J, Hampton CY, Hartmann A, Hartmann R, Hauriege S, Hauser G, Hirsemann H, Holl P, Holton JM, Homke A, Johansson L, Kimmel N, Kassemeyer S, Krasniqi F, Kuhnel K, Liang MN, Lomb L, Malmerberg E, Marchesini S, Martin AV, Maia FRNC, Messerschmidt M, Nass K, Reich C, Neutze R, Rolles D, Rudek B, Rudenko A, Schlichting I, Schmidt C,

Schmidt KE, Schulz J, Seibert MM, Shoeman RL, Sierra R, Soltau H, Starodub D, Stellato F, Stern S, Struder L, Timneanu N, Ullrich J, Wang XY, Williams GJ, Weidenspointner G, Weierstall U, Wunderer C, Barty A, Spence JCH, Chapman HN (2012) Time-resolved protein nanocrystallography using an X-ray free-electron laser. Optics Express 20: 2706-2716

Barber J (2009) Photosynthetic energy conversion: natural and artificial. Chem Soc Rev 38: 185-196

Barty A, Caleman C, Aquila A, Timneanu N, Lomb L, White TA, Andreasson J, Arnlund D, Bajt S, Barends TR, Barthelmess M, Bogan MJ, Bostedt C, Bozek JD, Coffee R, Coppola N, Davidsson J, Deponte DP, Doak RB, Ekeberg T, Elser V, Epp SW, Erk B, Fleckenstein H, Foucar L, Fromme P, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann R, Hartmann A, Hauser G, Hirsemann H, Holl P, Hunter MS, Johansson L, Kassemeyer S, Kimmel N, Kirian RA, Liang M, Maia FR, Malmerberg E, Marchesini S, Martin AV, Nass K, Neutze R, Reich C, Rolles D, Rudek B, Rudenko A, Scott H, Schlichting I, Schulz J, Seibert MM, Shoeman RL, Sierra RG, Soltau H, Spence JC, Stellato F, Stern S, Struder L, Ullrich J, Wang X, Weidenspointner G, Weierstall U, Wunderer CB, Chapman HN (2012) Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. Nat Photonics 6: 35-40

Barty A, Caleman C, Aquila A, Timneanu N, Lomb L, White TA, Andreasson J, Arnlund D, Bajt S, Barends TRM, Barthelmess M, Bogan MJ, Bostedt C, Bozek JD, Coffee R, Coppola N, Davidsson J, DePonte DP, Doak RB, Ekeberg T, Elser V, Epp SW, Erk B, Fleckenstein H, Foucar L, Fromme P, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann R, Hartmann A, Hauser G, Hirsemann H, Holl P, Hunter MS, Johansson L, Kassemeyer S, Kimmel N, Kirian RA, Liang MN, Maia FRNC, Malmerberg E, Marchesini S, Martin AV, Nass K, Neutze R, Reich C, Rolles D, Rudek B, Rudenko A, Scott H, Schlichting I, Schulz J, Seibert MM, Shoeman RL, Sierra RG, Soltau H, Spence JCH, Stellato F, Stern S, Struder L, Ullrich J, Wang X, Weidenspointner G, Weierstall U, Wunderer CB, Chapman HN (2012) Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. Nature Photonics 6: 35-40

Baxter RH, Ponomarenko N, Srajer V, Pahl R, Moffat K, Norris JR (2004) Time-resolved crystallographic studies of light-induced structural changes in the photosynthetic reaction center. Proc Natl Acad Sci U S A 101: 5982-5987

Bernal JD, Crowfoot D (1934) X-ray photographs of crystalline pepsin. Nature 133: 794-795

Bhat TN (1988) Calculation of an Omit Map. Journal of Applied Crystallography 21: 279-281

Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H (2011) Overcoming barriers to membrane protein structure determination. Nat Biotechnol 29: 335-340

Blake CC, Koenig DF, Mair GA, North AC, Phillips DC, Sarma VR (1965) Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. Nature 206: 757-761

Bodo G, Dintzis HM, Kendrew JC, Wyckoff HW (1959) The crystal structure of myoglobin: V.A. low-resolutin three-dimensional Fourier synthesis of sperm-whale myoglobin crystals. Proc. R. Soc. London Ser. A 253: 70-102

Bostedt C, Eremina E, Rupp D, Adolph M, Thomas H, Hoener M, de Castro AR, Tiggesbaumker J, Meiwes-Broer KH, Laarmann T, Wabnitz H, Plonjes E, Treusch R, Schneider JR, Moller T (2012) Ultrafast x-ray scattering of xenon nanoparticles: imaging transient states of matter. Phys Rev Lett 108: 093401

Bourgeois D, Vallone B, Arcovito A, Sciara G, Schotte F, Anfinrud PA, Brunori M (2006) Extended subnanosecond structural dynamics of myoglobin revealed by Laue crystallography. Proc Natl Acad Sci U S A 103: 4924-4929

Boutet S, Lomb L, Williams GJ, Barends TR, Aquila A, Doak RB, Weierstall U, DePonte DP, Steinbrener J, Shoeman RL, Messerschmidt M, Barty A, White TA, Kassemeyer S, Kirian RA, Seibert MM, Montanez PA, Kenney C, Herbst R, Hart P, Pines J, Haller G, Gruner SM, Philipp HT, Tate MW, Hromalik M, Koerner LJ, van Bakel N, Morse J, Ghonsalves W, Arnlund D, Bogan MJ, Caleman C, Fromme R, Hampton CY, Hunter MS, Johansson LC, Katona G, Kupitz C, Liang M, Martin AV, Nass K, Redecke L, Stellato F, Timneanu N, Wang D, Zatsepin NA, Schafer D, Defever J, Neutze R, Fromme P, Spence JC, Chapman HN, Schlichting I (2012) High-resolution protein structure determination by serial femtosecond crystallography. Science 337: 362-364

Boutet S, Lomb L, Williams GJ, Barends TRM, Aquila A, Doak RB, Weierstall U, DePonte DP, Steinbrener J, Shoeman RL, Messerschmidt M, Barty A, White TA, Kassemeyer S, Kirian RA, Seibert MM, Montanez PA, Kenney C, Herbst R, Hart P, Pines J, Haller G, Gruner SM, Philipp HT, Tate MW, Hromalik M, Koerner LJ, van Bakel N, Morse J, Ghonsalves W, Arnlund D, Bogan MJ, Caleman C, Fromme R, Hampton CY, Hunter MS, Johansson LC, Katona G, Kupitz C, Liang MN, Martin AV, Nass K, Redecke L, Stellato F, Timneanu N, Wang DJ, Zatsepin NA, Schafer D, Defever J, Neutze R, Fromme P, Spence JCH, Chapman HN, Schlichting I (2012) High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. Science 337: 362-364

Boutet S, Williams GJ (2010) The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). New Journal of Physics 12

Burmeister WP (2000) Structural changes in a cryo-cooled protein crystal owing to radiation damage. Acta Crystallogr D Biol Crystallogr 56: 328-341

Caffrey M (2009) Crystallizing membrane proteins for structure determination: use of lipidic mesophases. Annu Rev Biophys 38: 29-51

Caffrey M, Cherezov V (2009) Crystallizing membrane proteins using lipidic mesophases. Nat Protoc 4: 706-731

Ceccarelli EA, Arakaki AK, Cortez N, Carrillo N (2004) Functional plasticity and catalytic efficiency in plant and bacterial ferredoxin-NADP(H) reductases. Biochim Biophys Acta 1698: 155-165

Chapman HN, Barty A, Bogan MJ, Boutet S, Frank M, Hau-Riege S, Marchesini S, Woods BW, Bajt S, Benner WH, London RA, Plonjes E, Kuhlmann M, Treusch R, Dusterer S, Tschentscher T, Schneider JR, Spiller E, Moller T, Bostedt C, Hoener M, Shapiro D, Hodgson KO, Van der Spoel D, Burmeister F, Bergh M, Caleman C, Huldt G, Seibert M, Maia FR, Lee RW, Szoke A, Timneanu N, Hajdu J (2006) Femtosecond Diffractive Imaging with a Soft-X-ray Free-Electron Laser. Nature Physics 2: 839-843

Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FR, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang M, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Homke A, Reich C, Pietschner D, Struder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kuhnel KU, Messerschmidt M, Bozek JD, Hau-Riege SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Jonsson O, Svenda M, Stern S, Nass K, Andritschke R, Schroter CD, Krasniqi F, Bott M, Schmidt KE, Wang X, Grotjohann I, Holton JM, Barends TR, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Spence JC (2011) Femtosecond X-ray protein nanocrystallography. Nature 470: 73-77

Cherezov V, Fersi H, Caffrey M (2001) Crystallization screens: compatibility with the lipidic cubic phase for in meso crystallization of membrane proteins. Biophys J 81: 225-242

Chu HA, Nguyen AP, Debus RJ (1995) Amino acid residues that influence the binding of manganese or calcium to photosystem II. 1. The lumenal interhelical domains of the D1 polypeptide. Biochemistry 34: 5839-5858

Cox N, Pantazis DA, Neese F, Lubitz W (2013) Biological water oxidation. Acc Chem Res 46: 1588-1596

Dau H, Zaharieva I, Haumann M (2012) Recent developments in research on water oxidation by photosystem II. Curr Opin Chem Biol 16: 3-10

Davis KM, Kosheleva I, Henning RW, Seidler GT, Pushkar Y (2013) Kinetic Modeling of the X-ray-Induced Damage to a Metalloprotein. J Phys Chem B 117: 9161-9169

de Wijn R, van Gorkom HJ (2001) Kinetics of electron transfer from Q(a) to Q(b) in photosystem II. Biochemistry 40: 11912-11922

Debus RJ, Strickler MA, Walker LM, Hillier W (2005) No evidence from FTIR difference spectroscopy that aspartate-170 of the D1 polypeptide ligates a manganese ion that undergoes oxidation during the S0 to S1, S1 to S2, or S2 to S3 transitions in photosystem II. Biochemistry 44: 1367-1374

Dekker JP, Plijter JJ, Ouwehand L, Vangorkom HJ (1984) Kinetics of Manganese Redox Transitions in the Oxygen-Evolving Apparatus of Photosynthesis. Biochimica Et Biophysica Acta 767: 176-179

DePonte DP, Weierstall U, Schmidt K, Warner J, Starodub D, Spence JCH, Doak RB (2008) Gas dynamic virtual nozzle for generation of microscopic droplet streams. Journal of Physics D-Applied Physics 41

DePonte DP, Weierstall U, Starodub D, Schmidt K, Spence JC, Doak RB (2008) Gas Dynamic Virtual Nozzle for Generation of Microscopic Droplet Streams. Journal of Physics D: Applied Physics 41

Dilbeck PL, Hwang HJ, Zaharieva I, Gerencser L, Dau H, Burnap RL (2012) The D1-D61N mutation in Synechocystis sp. PCC 6803 allows the observation of pH-sensitive intermediates in the formation and release of O(2) from photosystem II. Biochemistry 51: 1079-1091

Dong X, Zhao Y, Huang X, Lin K, Chen J, Wei E, Liu T, Hu Y (2013) Structure-based drug design using GPCR homology modeling: toward the discovery of novel selective CysLT2 antagonists. Eur J Med Chem 62: 754-763

Emma P, Akre R, Arthur J, Bionta R, Bostedt C, Bozek J, Brachmann A, Bucksbaum P, Coffee R, Decker FJ, Ding Y, Dowell D, Edstrom S, Fisher A, Frisch J, Gilevich S, Hastings J, Hays G, Hering P, Huang Z, Iverson R, Loos H, Messerschmidt M, Miahnahri A, Moeller S, Nuhn HD, Pile G, Ratner D, Rzepiela J, Schultz D, Smith T, Stefan P, Tompkins H, Turner J, Welch J, White W, Wu J, Yocky G, Galayda J (2010) First lasing and operation of an angstrom-wavelength free-electron laser. Nature Photonics 4: 641-647

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66: 486-501

Fischer N, Hippler M, Setif P, Jacquot JP, Rochaix JD (1998) The PsaC subunit of photosystem I provides an essential lysine residue for fast electron transfer to ferredoxin. EMBO J 17: 849-858

Friedrich W, Knipping P, Von Laue M (1912) Proc. Bavarian Acad. Sci. : 303

Fromme P (1996) Structure and function of photosystem I. Curr Opin Struct Biol 6: 473-484

Fromme P, Grotjohann I (2008) Structure of Photosystems I and II. Results Probl Cell Differ 45: 33-72

Fromme P, Mathis P (2004) Unraveling the photosystem I reaction center: a history, or the sum of many efforts. Photosynth Res 80: 109-124

Fromme P, Melkozernov A, Jordan P, Krauss N (2003) Structure and function of photosystem I: interaction with its soluble electron carriers and external antenna systems. FEBS Lett 555: 40-44

Fromme P, Spence JC (2011) Femtosecond nanocrystallography using X-ray lasers for membrane protein structure determination. Curr Opin Struct Biol 21: 509-516

Fromme P, Witt HT (1998) Improved isolation and crystallization of Photosystem I for structural analysis. Biochim Biophys Acta 365: 175-184

Grundmeier A, Dau H (2012) Structural models of the manganese complex of photosystem II and mechanistic implications. Biochim Biophys Acta 1817: 88-105

Gust D, Moore TA, Moore AL (2009) Solar fuels via artificial photosynthesis. Acc Chem Res 42: 1890-1898

Haas DJ, Rossmann MG (1970) Crystallographic studies on lactate dehydrogenase at-75 degrees C. Acta Crystallogr B 26: 998-1004

Han G, Ho FM, Havelius KG, Morvaridi SF, Mamedov F, Styring S (2008) Direct quantification of the four individual S states in Photosystem II using EPR spectroscopy. Biochim Biophys Acta 1777: 496-503

Han G, Mamedov F, Styring S (2012) Misses during water oxidation in photosystem II are S state-dependent. J Biol Chem 287: 13422-13429

Han JH, Kang HS, Ko IS (2012) Status of the PAL-XFEL Project. Proceedings of the IPAC: 1735-1737

Hauptman HA (1991) History of X-ray crystallography. Chemometrics and Intelligent Laboratory Systems 10: 13

Herrmann S, Boutet S, Duda B, Fritz D, Haller G, Hart P, Herbst R, Kenney C, Lemke H, Messerschmidt M, Pines J, Robert A, Sikorski M, Williams G (2013) CSPAD-140k: A versatile detector for LCLS experiments. Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment 718: 550-553

Hodel A, Kim SH, Brunger AT (1992) Model Bias in Macromolecular Crystal-Structures. Acta Crystallographica Section A 48: 851-858

Holton JM (2009) A beginner's guide to radiation damage. J Synchrotron Radiat 16: 133-142

Hope H (1988) Cryocrystallography of biological macromolecules: a generally applicable method. Acta Crystallogr B 44 (Pt 1): 22-26

Howells MR, Beetz T, Chapman HN, Cui C, Holton JM, Jacobsen CJ, Kirz J, Lima E, Marchesini S, Miao H, Sayre D, Shapiro DA, Spence JC, Starodub D (2009) An assessment of the resolution limitation due to radiation-damage in x-ray diffraction microscopy. J Electron Spectros Relat Phenomena 170: 4-12

Hundelt M, Hays AM, Debus RJ, Junge W (1998) Oxygenic photosystem II: the mutation D1-D61N in Synechocystis sp. PCC 6803 retards S-state transitions without affecting electron transfer from YZ to P680+. Biochemistry 37: 14450-14456

Hunter MS, DePonte DP, Shapiro DA, Kirian RA, Wang X, Starodub D, Marchesini S, Weierstall U, Doak RB, Spence JC, Fromme P (2011) X-ray diffraction from membrane protein nanocrystals. Biophys J 100: 198-206

Hunter MS, Fromme P (2011) Toward structure determination using membrane-protein nanocrystals and microcrystals. Methods 55: 387-404

Institute P-S (2014) SwissFEL Accelerator Schematic. *In*,

Ishikawa T, Aoyagi H, Asaka T, Asano Y, Azumi N, Bizen T, Ego H, Fukami K, Fukui T, Yukito F, Goto S, Hanaki H, Hara T, Hasegawa T, Hatsui T, Higashiya A, Hirono T, Hosoda N, Ishii M, Inagaki T, Inubushi Y, Itoga T, Joti Y, Kago M, Kameshima T, Kimura H, Kiraihara Y, Kiyomichi A, Kobayashi T, Kondo C, Kudo T, Maesaka H, Marechal X, Masuda T, Matsubara S, Matsumoto T, Matsushita T, Matsui S, Nagasono M, Nariyama N, Ohashi H, Ohata T, Ohshima T, Ono S, Otake Y, Saji C, Sakurai T, Sato T, Sawada K, Seike T, Shirasawa K, Sugimoto T, Suzuki S, Takahashi S, Takebe H, Takeshita K, Tamasaku K, Tanaka H, Tanaka R, Tanaka T, Togashi T, Togawa K, Tokuhisa A, Tomizawa H, Tono K, Wu S, Yabashi M, Yamaga M, Yamashita A, Yanagida K, Zhang C,

Shintake T, Kitamura H, Kumagai N (2012) A compact X-ray free-electron laser emitting in the sub-angstrom region. Nature Photonics 6: 540-544

Isobe H, Shoji M, Yamanaka S, Umena Y, Kawakami K, Kamiya N, Shen JR, Yamaguchi K (2012) Theoretical illumination of water-inserted structures of the CaMn4O5 cluster in the S2 and S3 states of oxygen-evolving complex of photosystem II: full geometry optimizations by B3LYP hybrid density functional. Dalton Trans 41: 13727-13740

Johnson LN, Phillips DC (1965) Structure of some crystalline lysozyme-inhibitor complexes determined by X-ray analysis at 6 Angstrom resolution. Nature 206: 761-763

Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution. Nature 411: 909-917

Kaminski R, Graber T, Benedict JB, Henning R, Chen YS, Scheins S, Messerschmidt M, Coppens P (2010) Optimizing the accuracy and precision of the single-pulse Laue technique for synchrotron photo-crystallography. J Synchrotron Radiat 17: 479-485

Kanady JS, Tsui EY, Day MW, Agapie T (2011) A synthetic model of the Mn(3)Ca subsite of the oxygen-evolving complex in photosystem II. Science 333: 733-736

Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. Science 336: 1030-1033

Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 181: 662-666

Kern J, Alonso-Mori R, Hellmich J, Tran R, Hattne J, Laksmono H, Glockner C, Echols N, Sierra RG, Sellberg J, Lassalle-Kaiser B, Gildea RJ, Glatzel P, Grosse-Kunstleve RW, Latimer MJ, McQueen TA, DiFiore D, Fry AR, Messerschmidt M, Miahnahri A, Schafer DW, Seibert MM, Sokaras D, Weng TC, Zwart PH, White WE, Adams PD, Bogan MJ, Boutet S, Williams GJ, Messinger J, Sauter NK, Zouni A, Bergmann U, Yano J, Yachandra VK (2012) Room temperature femtosecond X-ray diffraction of photosystem II microcrystals. Proc Natl Acad Sci U S A 109: 9721-9726

Kern J, Alonso-Mori R, Tran R, Hattne J, Gildea RJ, Echols N, Glockner C, Hellmich J, Laksmono H, Sierra RG, Lassalle-Kaiser B, Koroidov S, Lampe A, Han G, Gul S, Difiore D, Milathianaki D, Fry AR, Miahnahri A, Schafer DW, Messerschmidt M, Seibert MM, Koglin JE, Sokaras D, Weng TC, Sellberg J, Latimer MJ, Grosse-Kunstleve RW, Zwart PH, White WE, Glatzel P, Adams PD, Bogan MJ,

Williams GJ, Boutet S, Messinger J, Zouni A, Sauter NK, Yachandra VK, Bergmann U, Yano J (2013) Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. Science 340: 491-495

Kern J, Loll B, Luneberg C, DiFiore D, Biesiadka J, Irrgang KD, Zouni A (2005) Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. Biochim Biophys Acta 1706: 147-157

Kirian RA, Wang X, Weierstall U, Schmidt KE, Spence JC, Hunter M, Fromme P, White T, Chapman HN, Holton J (2010) Femtosecond protein nanocrystallography-data analysis methods. Opt Express 18: 5713-5723

Kirian RA, White TA, Holton JM, Chapman HN, Fromme P, Barty A, Lomb L, Aquila A, Maia FR, Martin AV, Fromme R, Wang X, Hunter MS, Schmidt KE, Spence JC (2011) Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals. Acta Crystallogr A 67: 131-140

Klukas O, Schubert WD, Jordan P, Krauss N, Fromme P, Witt HT, Saenger W (1999) Photosystem I, an improved model of the stromal subunits PsaC, PsaD, and PsaE. J Biol Chem 274: 7351-7360

Kohoutova J, Kuta Smatanova I, Brynda J, Lapkouski M, Revuelta JL, Arellano JB, Ettrich R (2009) Crystallization and preliminary crystallographic characterization of the extrinsic PsbP protein of photosystem II from Spinacia oleracea. Acta Crystallogr Sect F Struct Biol Cryst Commun 65: 111-115

Kok B, Forbush B, McGloin M (1970) Cooperation of charges in photosynthetic O2 evolution-I. A linear four step mechanism. Photochem Photobiol 11: 457-475

Krauss N, Hinrichs W, Witt I, Fromme P, Pritzkow W, Dauter Z, Betzel C, Wilson KS, Witt HT, Saenger W (1993) 3-dimensional structure of system-1 of photosynthesis at 6 angstrom resolution. nature 361: 326-331

Krauss N, Schubert WD, Klukas O, Fromme P, Witt HT, Saenger W (1996) Photosystem I at 4 A resolution represents the first structural model of a joint photosynthetic reaction centre and core antenna system. Nat Struct Biol 3: 965-973

Kupitz C, Basu S, Grotjohann I, Fromme R, Zatsepin NA, Rendek KN, Hunter M, Shoeman RL, White TA, Wang D, James D, Yang J, Cobb DE, Reeder B, Sierra RG, Liu H, Barty A, Aquila AL, Deponte D, Kirian RA, Bari S, Bergkamp JJ, Beyerlein KR, Bogan MJ, Caleman C, Chao T, Conrad CE, Davis KM, Fleckenstein H, Galli L, Hau-Riege SP, Kassemeyer S, Laksmono H, Liang M, Lomb L, Marchesini S, Martin AV, Messerschmidt M, Milathianaki D, Nass K, Ros A, Roy-Chowdhury S, Schmidt K, Seibert M, Steinbrener J, Stellato F, Yan L, Yoon C, Moore TA, Moore AL, Pushkar Y, Williams GJ, Boutet S, Doak RB, Weierstall U, Frank M, Chapman HN, Spence JCH, Fromme P (2014) Serial

time-resolved crystallography of photosystem II using a femtosecond X-ray laser. Nature

Kurisu G, Zhang H, Smith JL, Cramer WA (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. Science 302: 1009-1014

Lee H, Cherni I, Yu H, Fromme R, Doran JD, Grotjohann I, Mittman M, Basu S, Deb A, Dorner KH, Aquila A, Barty A, Boutet S, Chapman HN, Doak RB, Hunter MS, James DR, Kirian RA, Kupitz C, Seibert M, Shoeman RL, Spence JC, Stellato F, Weierstall U, Williams GJ, Yoon C, Wang D, Zatsepin NA, Hogue BG, Matoba N, Fromme P, Mor TS (2014) Expression, purification and crystallization of CTB-MPR, a candidate mucosal vaccine component against HIV-1. IUCrJ In Publication

Liu W, Wacker D, Gati C, Han GW, James D, Wang D, Nelson G, Weierstall U, Katritch V, Barty A, Zatsepin NA, Li D, Messerschmidt M, Boutet S, Williams GJ, Koglin JE, Seibert MM, Wang C, Shah ST, Basu S, Fromme R, Kupitz C, Rendek KN, Grotjohann I, Fromme P, Kirian RA, Beyerlein KR, White TA, Chapman HN, Caffrey M, Spence JC, Stevens RC, Cherezov V (2013) Serial femtosecond crystallography of G protein-coupled receptors. Science 342: 1521-1524

Loll B, Kern J, Saenger W, Zouni A, Biesiadka J (2005) Towards complete cofactor arrangement in the 3.0 A resolution structure of photosystem II. Nature 438: 1040-1044

Lomb L, Steinbrener J, Bari S, Beisel D, Berndt D, Kieser C, Lukat M, Neef N, Shoeman RL (2012) An anti-settling sample delivery instrument for serial femtosecond crystallography. Journal of Applied Crystallography 45: 674-678

Luber S, Rivalta I, Umena Y, Kawakami K, Shen JR, Kamiya N, Brudvig GW, Batista VS (2011) S1-state model of the O2-evolving complex of photosystem II. Biochemistry 50: 6308-6311

Marchesini S, He H, Chapman HN, Hau-Riege S, Noy A, Howells MR, Weierstall U, Spence JC (2003) X-ray Image Reconstruction from a Diffraction Pattern Alone. Phys. Rev. B. 68: 140101

McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr 63: 32-41

Meents A, Gutmann S, Wagner A, Schulze-Briese C (2010) Origin and temperature dependence of radiation damage in biological samples at cryogenic temperatures. Proc Natl Acad Sci U S A 107: 1094-1099

Moraes I, Evans G, Sanchez-Weatherby J, Newstead S, Stewart PD (2014) Membrane protein structure determination - the next generation. Biochim Biophys Acta 1838: 78-87

Muller MG, Slavov C, Luthra R, Redding KE, Holzwarth AR (2010) Independent initiation of primary electron transfer in the two branches of the photosystem I reaction center. Proc Natl Acad Sci U S A 107: 4123-4128

Navarro MP, Ames WM, Nilsson H, Lohmiller T, Pantazis DA, Rapatskiy L, Nowaczyk MM, Neese F, Boussac A, Messinger J, Lubitz W, Cox N (2013) Ammonia binding to the oxygen-evolving complex of photosystem II identifies the solvent-exchangeable Q:1; 2; 3 μ-oxo of the manganese tetramer. PNAS DOI 10.1073

Nederlof I, Li YW, van Heel M, Abrahams JP (2013) Imaging protein three-dimensional nanocrystals with cryo-EM. Acta Crystallogr D Biol Crystallogr 69: 852-859

Neutze R, Moffat K (2012) Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges. Curr Opin Struct Biol 22: 651-659

Neutze R, Wouts R, van der Spoel D, Weckert E, Hajdu J (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. Nature 406: 752-757

Philipp HT, Hromalik M, Tate M, Koerner L, Gruner SM (2011) Pixel array detector for X-ray free electron laser experiments. Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment 649: 67-69

Pushkar Y, Yano J, Sauer K, Boussac A, Yachandra VK (2008) Structural changes in the Mn4Ca cluster and the mechanism of photosynthetic water splitting. Proc Natl Acad Sci U S A 105: 1879-1884

Rajagopal S, Anderson S, Srajer V, Schmidt M, Pahl R, Moffat K (2005) A structural pathway for signaling in the E46Q mutant of photoactive yellow protein. Structure 13: 55-63

Read RJ (2003) New ways of looking at experimental phasing. Acta Crystallographica Section D-Biological Crystallography 59: 1891-1902

Redecke L, Nass K, DePonte DP, White TA, Rehders D, Barty A, Stellato F, Liang M, Barends TR, Boutet S, Williams GJ, Messerschmidt M, Seibert MM, Aquila A, Arnlund D, Bajt S, Barth T, Bogan MJ, Caleman C, Chao TC, Doak RB, Fleckenstein H, Frank M, Fromme R, Galli L, Grotjohann I, Hunter MS, Johansson LC, Kassemeyer S, Katona G, Kirian RA, Koopmann R, Kupitz C, Lomb L, Martin AV, Mogk S, Neutze R, Shoeman RL, Steinbrener J, Timneanu N, Wang D, Weierstall U, Zatsepin NA, Spence JC, Fromme P, Schlichting I, Duszenko M, Betzel C, Chapman HN (2013) Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser. Science 339: 227-230

Renger G (2012) Mechanism of light induced water splitting in Photosystem II of oxygen evolving photosynthetic organisms. Biochim Biophys Acta 1817: 1164-1176

Rivalta I, Brudvig GW, Batista VS (2012) Oxomanganese complexes for natural and artificial photosynthesis. Curr Opin Chem Biol 16: 11-18

Rowe EM, Mills FE (1973) Tantaulus I: A Dedicated Storage Ring Synchrotron Radiation Source. Particle Accelerators 4: 211-227

Sauter NK, Hattne J, Grosse-Kunstleve RW, Echols N (2013) New Python-based methods for data processing. Acta Crystallographica Section D-Biological Crystallography 69: 1274-1282

Schmidt M, Graber T, Henning R, Srajer V (2010) Five-dimensional crystallography. Acta Crystallogr A 66: 198-206

Schmidt M, Srajer V, Purwar N, Tripathi S (2012) The kinetic dose limit in room-temperature time-resolved macromolecular crystallography. Journal of Synchrotron Radiation 19: 264-273

Setif PQ, Bottin H (1995) Laser flash absorption spectroscopy study of ferredoxin reduction by photosystem I: spectral and kinetic evidence for the existence of several photosystem I-ferredoxin complexes. Biochemistry 34: 9059-9070

Shen JR, Kawakami K, Koike H (2011) Purification and crystallization of oxygen-evolving photosystem II core complex from thermophilic cyanobacteria. Methods Mol Biol 684: 41-51

Shinkarev VP (2003) Oxygen evolution in photosynthesis: simple analytical solution for the Kok model. Biophys J 85: 435-441

Siegbahn PE (2008) Theoretical studies of O-O bond formation in photosystem II. Inorg Chem 47: 1779-1786

Siegbahn PE (2012) Water oxidation mechanism in photosystem II, including oxidations, proton release pathways, O-O bond formation and O2 release. Biochim Biophys Acta 1827: 1003-1019

Siegbahn PE (2013) Substrate water exchange for the oxygen evolving complex in PSII in the S1, S2, and S3 states. J Am Chem Soc 135: 9442-9449

Sierra RG, Laksmono H, Kern J, Tran R, Hattne J, Alonso-Mori R, Lassalle-Kaiser B, Glockner C, Hellmich J, Schafer DW, Echols N, Gildea RJ, Grosse-Kunstleve RW, Sellberg J, McQueen TA, Fry AR, Messerschmidt MM, Miahnahri A, Seibert MM, Hampton CY, Starodub D, Loh ND, Sokaras D, Weng TC, Zwart PH, Glatzel P, Milathianaki D, White WE, Adams PD, Williams GJ, Boutet S,

Zouni A, Messinger J, Sauter NK, Bergmann U, Yano J, Yachandra VK, Bogan MJ (2012) Nanoflow electrospinning serial femtosecond crystallography. Acta Crystallogr D Biol Crystallogr 68: 1584-1587

Spence JC, Kirian R, Wang X, Weierstall U, Schimdt KE, White T, Barty A, Chapman HN, Marchesini S, Holton J (2011) Phasing of Coherent Femtosecond X-ray Diffractoin from Size-Varying Nanocrystals. Opt. Express 19: 2866-2873

Spence JC, Weierstall U, Chapman HN (2012) X-ray lasers for structural and dynamic biology. Rep Prog Phys 75: 102601

Spence JCH, Weierstall U, Chapman HN (2012) X-ray lasers for structural and dynamic biology. Reports on Progress in Physics 75

Takahashi S, Badger MR (2011) Photoprotection in plants: a new light on photosystem II damage. Trends Plant Sci 16: 53-60

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Adams PD, Read RJ, Zwart PH, Hung LW (2008) Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias. Acta Crystallographica Section D-Biological Crystallography 64: 515-524

Tripathi S, Srajer V, Purwar N, Henning R, Schmidt M (2012) pH dependence of the photoactive yellow protein photocycle investigated by time-resolved crystallography. Biophys J 102: 325-332

Umena Y, Kawakami K, Shen JR, Kamiya N (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A. Nature 473: 55-60

Vellieux FMD, Dijkstra BW (1997) Computation of Bhat's OMIT maps with different coefficients. Journal of Applied Crystallography 30: 396-399

Wampler RD, Kissick DJ, Dehen CJ, Gualtieri EJ, Grey JL, Wang HF, Thompson DH, Cheng JX, Simpson GJ (2008) Selective detection of protein crystals by second harmonic microscopy. J Am Chem Soc 130: 14076-14077

Weierstall U, Doak RB, Spence JC, Starodub D, Shapiro D, Kennedy P, Warner J, Hembree GG, Fromme P, Chapman HN (2008) Droplet streams for serial crystallography of proteins. Exp. Fluids 44: 675-689

Weierstall U, Doak RB, Spence JCH, Starodub D, Shapiro D, Kennedy P, Warner J, Hembree GG, Fromme P, Chapman HN (2008) Droplet streams for serial crystallography of proteins. Experiments in Fluids 44: 675-689

Weierstall U, James D, Wang C, White T, Wang D, Liu W, Spence J, Doak R, Nelson G, Fromme P, Fromme R, Grotjohann I, Kupitz C, Zatsepin N, Liu H, Basu S,

Wacker D, Han G, Katritch V, Boutet S, Messerschmidt M, Williams G, Koglin J, Seibert M, Klinker M, Gati C, Shoeman R, Barty A, Chapman H, Kirian R, Beyerlein K, Stevens R, Li D, Shah S, Howe N, Caffrey M, Cherezov V (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. Nat Commun 5: 3309

Weierstall U, James D, Wang C, White TA, Wang D, Liu W, Spence JC, Bruce Doak R, Nelson G, Fromme P, Fromme R, Grotjohann I, Kupitz C, Zatsepin NA, Liu H, Basu S, Wacker D, Han GW, Katritch V, Boutet S, Messerschmidt M, Williams GJ, Koglin JE, Marvin Seibert M, Klinker M, Gati C, Shoeman RL, Barty A, Chapman HN, Kirian RA, Beyerlein KR, Stevens RC, Li D, Shah ST, Howe N, Caffrey M, Cherezov V (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. Nat Commun 5: 3309

Weierstall U, Spence JC, Doak RB (2012) Injector for scattering measurements on fully solvated biospecies. Rev Sci Instrum 83: 035108

Weierstall U, Spence JCH, Doak RB (2012) Injector for scattering measurements on fully solvated biospecies. Review of Scientific Instruments 83

White TA, Kirian RA, Martin AV, Aquila A, Nass K, Barty A, Chapman HN (2012) CrystFEL: a software suite for snapshot serial crystallography. Journal of Applied Crystallography 45: 335-341

White TA, Kirian RA, Martin AV, Aquila A, Nass K, Barty A, Chapman HN (2012) CrystFEL: a software suite for snapshot serial crystallography. Journal of Applied Crystallography 45: 335-341

Witt HT, Witt I, Difiore D, Rogner M, Hinrichs W, Saenger W, Granzin J, Betzel C, Dauter Z (1988) X-ray characterization of single-crystals of the reaction center-I of water splitting photosynthesis. Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics 92: 1503-1506

Witt HT, Witt I, Krauss N, Hinrichs W, Fromme P, Saenger W (1994) Crystals and structure of Photosystem-I. Biophysical Journal 66: A2-A2

Witt I, Witt HT, Difiore D, Rogner M, Hinrichs W, Saenger W, Granzin J, Betzel C, Dauter Z (1988) X-ray characterization of single-crystals of the reaction center-I of water splitting photosynthesis. Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics 92: 1503-1506

Woolfson MM (1997) An introduction to X-ray crystallography, Ed 2nd. Cambridge University Press, Cambridge ; New York, NY, USA

Yamaguchi K, Yamanaka S, Isobe H, Saito T, Kanda K, Umena Y, Kawakami K, Shen JR, Kamiya N, Okumura M, Nakamura H, Shoji M, Yoshioka Y (2013) The nature of chemical bonds of the CaMn4O5 cluster in oxygen evolving complex of

photosystem II: Jahn-Teller distortion and its suppression by Ca doping in cubane structures. International Journal of Quantum Chemistry 113: 453-473

Yano J, Kern J, Irrgang KD, Latimer MJ, Bergmann U, Glatzel P, Pushkar Y, Biesiadka J, Loll B, Sauer K, Messinger J, Zouni A, Yachandra VK (2005) X-ray damage to the Mn4Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography. Proc Natl Acad Sci U S A 102: 12047-12052

Yano J, Kern J, Pushkar Y, Sauer K, Glatzel P, Bergmann U, Messinger J, Zouni A, Yachandra VK (2008) High-resolution structure of the photosynthetic Mn4Ca catalyst from X-ray spectroscopy. Philos Trans R Soc Lond B Biol Sci 363: 1139-1147; discussion 1147

Yano J, Yachandra VK (2007) Oxidation state changes of the Mn4Ca cluster in photosystem II. Photosynth Res 92: 289-303

Yano J, Yachandra VK (2008) Where water is oxidized to dioxygen: structure of the photosynthetic Mn4Ca cluster from X-ray spectroscopy. Inorg Chem 47: 1711-1726

Zastrau U, Sperling P, Harmand M, Becker A, Bornath T, Bredow R, Dziarzhytski S, Fennel T, Fletcher LB, Forster E, Gode S, Gregori G, Hilbert V, Hochhaus D, Holst B, Laarmann T, Lee HJ, Ma T, Mithen JP, Mitzner R, Murphy CD, Nakatsutsumi M, Neumayer P, Przystawik A, Roling S, Schulz M, Siemer B, Skruszewicz S, Tiggesbaumker J, Toleikis S, Tschentscher T, White T, Wostmann M, Zacharias H, Doppner T, Glenzer SH, Redmer R (2014) Resolving ultrafast heating of dense cryogenic hydrogen. Phys Rev Lett 112: 105002

Zouni A, Jordan R, Schlodder E, Fromme P, Witt HT (2000) First photosystem II crystals capable of water oxidation. Biochim Biophys Acta 1457: 103-105

Zouni A, Witt HT, Kern J, Fromme P, Krauss N, Saenger W, Orth P (2001) Crystal structure of photosystem II from Synechococcus elongatus at 3.8 A resolution. Nature 409: 739-743

APPENDIX A

TABLES OF DATA SET STATISTICS

Table A.I. Statistics of the Dark (S_1) and Double-Flash (Putative S_3) Data Sets Collected at CXI, LCLS

	Dark data set	Double-flash data set
Wavelength (Å)	2.05	2.05
Resolution range (Å)	100.6 - 5.0	102.3 - 5.5
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit Cell length (Å)	133.3 ± 1.6, 226.3 ± 2.1, 307.1 ± 3.1	136.6 ± 1.5, 228.1 ± 2.3, 308.7 ± 3.8
Total reflections	28,679,554 (1,679,683)	12,476,013 (1,018,721)
Unique reflections	40,946 (2,710)	32,066 (2,651)
Multiplicity	700.35 (618.0)	388.55 (381.1)
Completeness (%)	99.98 (100)	99.88 (100)
Mean I/σ(I)	10.65 (2.1)	8.03 (1.75)
$CC_{1/2}$ ‡	0.914 (0.740)	0.877 (0.635)
R_{split}	0.07 (0.37)	0.09 (0.49)
R_{work}	0.260 (0.3502)	0.280 (0.3820)
R_{free}	0.262 (0.3434)	0.290 (0.3477)
RMS* (bonds) Å	0.039	0.039
RMS* (angles) deg	3.029	3.029
Number of atoms	49,817	49,817
Protein residues	5,214	5,214
Ramachandran* favored (%)	97.7	97.7
Ramachandran* outliers (%)	0.2	0.2
Clashscore (Molprobity)	5.5	5.8
Mean B-factor# ($Å^2$)	33.7	33.7

Statistics of the dark (S_1) and double-flash (putative S_3) data sets collected at CXI, LCLS. Numbers in parentheses refer to values for the highest resolution shells.

** $CC_{1/2}$ is Pearson's coefficient calculated as described in Karplus et al. (Karplus and Diederichs, 2012)*
**It is noted that the values for the RMS for bonds and angles as well as the Ramachandran values are positively biased by the high resolution model 3ARC.*
 # the B-factors were taken from the high resolution model 3ARC and not refined.

$$R_{split} = \sqrt{2}\frac{\sum_{hkl} | I_{even} - I_{odd} |}{\sum_{hkl} | I_{even} + I_{odd} |}$$

Table A.II. Statistics of the Femtosecond Crystallography X-ray Diffraction Data Sets

a

Resolution (Å)	observed unique reflections	Measurements	Completeness (%)	Multiplicity	I/σ(I)	$CC_{1/2}$	Reflections (used for $CC_{1/2}$)
19.20	4,255	3,081,618	99.63	724.2	23.12	0.990	4,236
9.60	4,038	3,048,418	99.48	754.9	16.69	0.991	4,017
8.03	3,980	3,064,098	99.20	769.9	12.26	0.989	3,926
7.17	3,928	2,847,009	98.35	724.8	8.04	0.977	3,783
6.58	3,826	2,608,679	96.28	681.8	5.67	0.957	3,552
6.16	3,808	2,497,532	95.70	655.9	4.31	0.936	3,446
5.82	3,682	2,552,117	92.91	693.1	3.59	0.903	3,200
5.55	3,663	2,548,799	92.73	695.8	3.21	0.890	3,105
5.32	3,543	2,489,203	90.04	702.6	2.75	0.853	2,822
5.13	3,426	2,241,076	86.95	654.1	2.41	0.826	2,664
4.96	3,555	2,193,921	90.46	617.1	2.29	0.740	2,767
4.81	3,518	2,241,551	89.84	637.2	2.10	0.695	2,655
4.68	3,448	2,246,609	88.23	651.6	1.89	0.392	2,549
4.56	3,439	2,201,217	86.95	640.1	1.77	0.463	2,442
4.45	3,316	2,176,673	84.85	656.4	1.69	0.445	2,327
4.36	3,166	2,097,792	81.87	662.6	1.52	0.216	2,078
4.27	3,116	2,111,675	78.86	677.7	1.47	0.135	1,963
4.18	2,995	2,097,219	77.03	700.2	1.40	0.070	1,841
4.11	2,850	2,009,233	73.00	705.0	1.27	0.062	1,607
4.03	2,930	2,032,330	74.99	693.6	1.32	0.068	1,721
Overall	**70,482**	**48,386,769**	**88.85**	**684.93**	**5.42**	**0.994**	**56,691**

b

Resolution (Å)	observed unique reflections	Measurements	Completeness (%)	Multiplicity	I/σ(I)	$CC_{1/2}$	Reflections (used for $CC_{1/2}$)
19.20	4,395	1,724,037	99.30	392.3	16.30	0.974	4,361
9.60	4,157	1,728,443	98.79	415.8	11.84	0.983	4.103
8.03	4,085	1,718,604	97.80	420.7	8.71	0.977	3,960
7.17	3,916	1,546,207	94.73	394.8	5.72	0.963	3,628
6.58	3,747	1,389,049	90.73	370.7	3.93	0.926	3,242
6.16	3,671	1,308,516	88.97	356.4	2.91	0.860	3,001
5.82	3,463	1,306,601	84.28	377.3	2.35	0.789	2,626
5.55	3,365	1,280,059	82.27	380.4	2.06	0.785	2,439
5.32	3,193	1,221,532	77.90	382.6	1.79	0.635	2,202
5.13	2,946	1,046,614	72.05	355.3	1.60	0.474	1,893
4.96	3,107	1,040,022	76.28	334.7	1.51	0.450	1,935
4.81	3,046	1,056,445	74.84	346.8	1.45	0.308	1,847
4.68	3,078	1,087,015	75.20	353.2	1.40	0.167	1,890
4.56	2,968	1,034,930	73.16	348.7	1.34	0.159	1,764
4.45	2,859	1,021,562	70.26	357.3	1.30	0.086	1,649
4.36	2,771	1,003,272	68.20	362.1	1.22	0.007	1,510
4.27	2,652	979,252	65.37	369.3	1.17	0.042	1,394
4.18	2,628	1,001,005	65.15	380.9	1.18	0.072	1,386
4.11	2,447	937,143	59.95	383.0	1.13	0.010	1,294
4.03	2,579	976,398	64.00	378.6	1.14	0.037	1,350
Overall	**65,073**	**24,496,706**	**78.96**	**373.04**	**4.09**	**0.988**	**47,474**

(a) of the dark (S_1) state by resolution bins*. **(b)** Statistics of the femtosecond crystallography X-ray diffraction data set of the double-flash (putative S_3) state by resolution bins.

* $CC_{1/2}$ is Pearson's coefficient calculated as described in Karplus et al.(Karplus and Diederichs, 2012).

Table A.III. Data Statistic Comparison of Hit and Indexing Rates as well as the Unit Cell Constants from Four Different Data Sets Collected on PSII Crystals.

a

	Dark (A)	2-flash(A)	Dark (B)	3-flash(B)
Crystal hits	71,628	63,363	33,373	32,190
Indexed images	34,554	18,772	11,664	12,543
Indexing rate (%)	48	29	35	38
Unit cell constant a (Å)	133.3 ± 1.6	136.6 ± 1.5	132.3 ± 1.2	132.3 ± 1.1
Unit cell constant b (Å)	226.3 ± 2.1	228.1 ± 2.3	226.1 ± 1.7	226.1 ± 1.6
Unit cell constant c (Å)	307.1 ± 3.1	308.7 ± 3.8	307.3 ± 2.8	307.0 ± 2.7

b

	Dark (A) (only dark)	Dark (A) (dark alternating with illumination)
Crystal hits	12,087	59,541
Indexed images	5,170	29,384
Indexing rate (%)	42.77	49.35
Unit cell constant a (Å)	133.2 ± 1.9	133.3 ± 1.5
Unit cell constant b (Å)	225.8 ± 2.2	226.3 ± 2.0
Unit cell constant c (Å)	307.6 ± 3.2	307.1 ± 3.1

c

	Kern et. al (2013)[18]		This work	
	S_1	S_2	Dark (S_1)	2-flash (putative S_3)
Resolution range (Å)	82.9 - 5.7 (5.9 - 5.7)	83.0 - 5.9 (6.1 - 5.9)	100.6 - 5.0 (5.11 - 5.0)	102.3 - 5.5 (5.66 - 5.5)
Total reflections	1,475,630 (8,036)	564,722 (8,662)	28,679,554 (1,679,683)	12,476,013 (1,018,721)
Unique reflections	27,220 (2,290)	24,671 (2,143)	40,946 (2,710)	32,066 (2,651)
Multiplicity	54.2 (3.5)	22.9 (4.0)	700.35 (618.0)	388.55 (381.1)
$CC_{1/2}$	0.802 (0.343)	0.661 (0.376)	0.914 (0.740)	0.877 (0.635)
R_{work}	0.276 (0.360)	0.284 (0.306)	0.260 (0.350)	0.280 (0.382)
R_{free}	0.315 (0.350)	0.317 (0.387)	0.262 (0.343)	0.290 (0.347)

(a) The data sets dark (A) and double-flash (A) were collected at the CXI instrument in January 2012 and may represent the dark S_1-state and putative S_3-state of PSII for which data evaluation and structural changes are discussed in this work. For comparison, the statistics are shown for data sets collected on the dark state S_1 and the transient triple-flash state (i.e., putative S_4-state) at the CXI instrument in June 2012.
(b) Data statistics for dark (A) separated into runs where the laser was switched off (only dark state) and dark state images from runs where alternate dark and light states patterns were recorded**. (c)** Data statistics from this work and from Kern et al. 2013 (Kern et al., 2013)**.** Comparison of the data evaluation statistics of the dark S_1-state and double-flash (putative S_3) state data from this work evaluated by the CrystFEL software suite (White et al., 2012) along with data from Kern et al. 2013 (Kern et al., 2013) on the dark S_1-state and the single excited S_2 -state evaluated with the software suite cctbx.xfel (Sauter et al., 2013). The numbers in brackets refer to the data statistics in the highest resolution shell.

APPENDIX B

SUPPORTING ELECTRON DENSITY MAPS

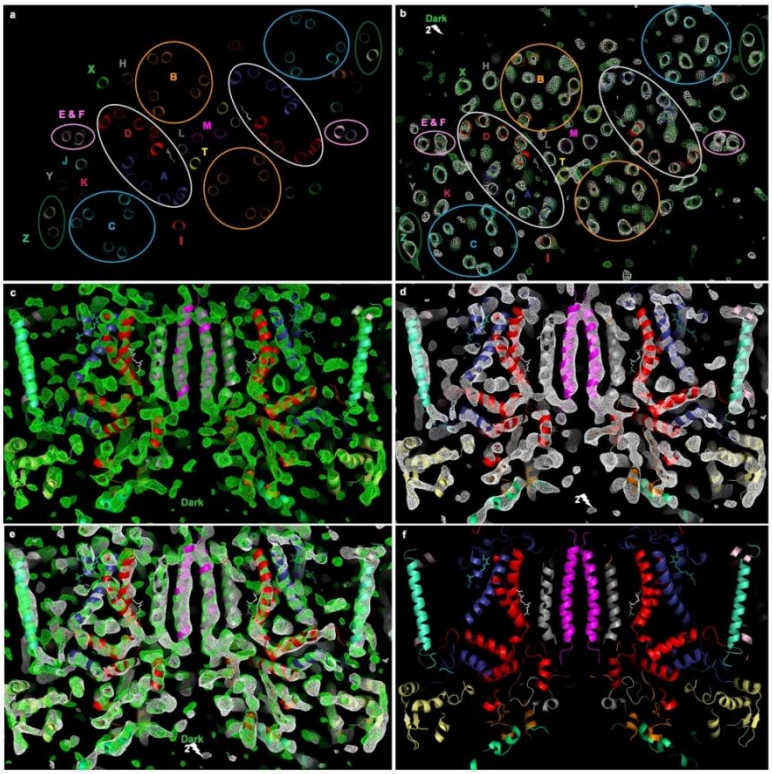


Figure A.1. The Arrangement of the Transmembrane Helices in the PSII Dimer. ***(a-b)*** *panels show an overview of the arrangement of transmembrane helices in Photosystem II. The protein subunits are named according to their genes so PsbA is subunit A, PsbB is subunit B, etc. The identification of the location of subunits with more than one transmembrane helix is facilitated by ovals, which are labeled using the same color code as the corresponding subunit.* ***(a)*** *Top view from the stromal side of the arrangement of transmembrane helices in the middle of the membrane. The assignment of helices to different protein subunits is based on the structural assignments of Umena et al.(Umena et al., 2011) The 5 transmembrane helices of the core subunits of the reaction center are PsbA (blue) and PsbB (red).* ***(b)*** *The picture shows the omit maps ($2F_o$-F_c) of the dark and double-flash states at the contour level of 1.5 σ in the same view direction as shown in* ***(a)***. ***(c-f)*** *These panels show that most of the alpha helices in the middle of the membrane are well matched between the dark and double-flash states, in the reaction center core (PsbA and PsbB) as well as in the peripheral parts of PSII (for example PsbZ). The view and color coding of helices are the same as in* ***(a)***. ***(c, d)*** *Omit maps of the dark (green) and double-flash (white) states of PSII show a cut through the central region of PSII at 1.5 σ.* ***(e)*** *The superposition of omit maps indicates a good general overlay of the transmembrane helices and the lumenal loop regions in the two omit maps featuring the reaction center core subunits PsbA (blue) and PsbB (red) as well as the peripheral subunit X (cyan), and the subunits M (pink) and L (grey) in the dimerization domain of the PSII dimer. The electron density is shown at the contour level of 1.5 σ.* ***(f)*** *The structural model is also shown with same color codes as in* ***(a)***.

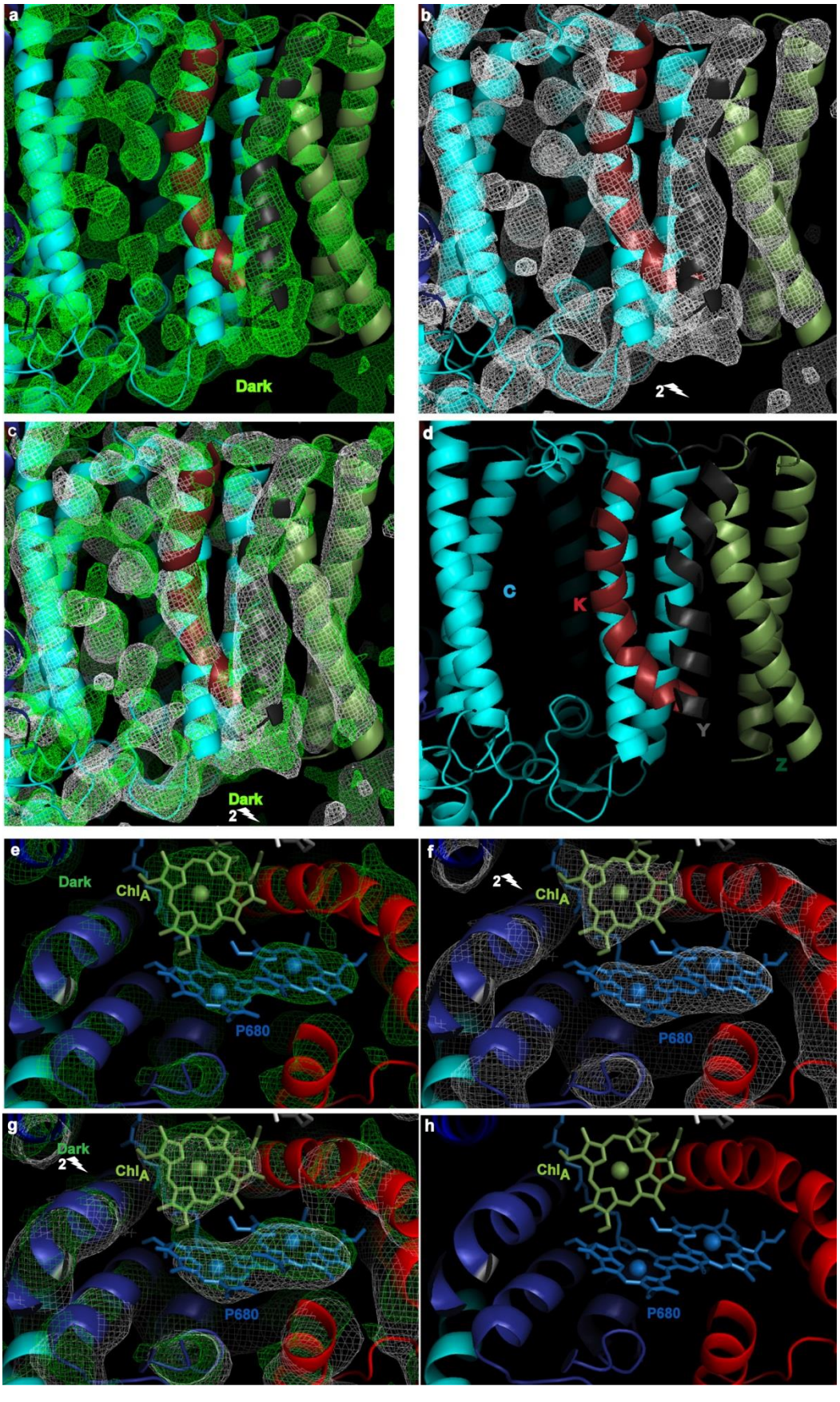

a
Dark
b
2
c
Dark
2
d
C
K
Y
Z
e
Dark
Chl A
P680
f
2
Chl A
P680
g
Dark
2
Chl A
P680
h
Chl A
P680

Figure A.2. Omit Map of the Dark and Double-Flash States of Several Important PSII Sites. *(**a-d**) This picture features the peripheral subunits PsbZ (grey-green), PsbK (brown), PsaH (grey) and the core-antenna protein CP43 (PsbC) (cyan). The omit electron density map at the contour level of 1.5 σ for the dark (S_1) state is depicted in green (**a**) and the double flash (putative S_3) state is depicted in white (**b**). The overlay of the two omit maps (**c**) is shown at the contour level of 1.5 σ. The globular densities between PsbC and PsbK correspond to antenna chlorophylls. The figure shows that even the electron density for the two most peripheral helices that belong to subunit PsbZ are well defined. We also note the good match of the strongly kinked helix of PsbK between the S_1 and S_3-state electron density maps. The subunits are labeled according to their genes in the view of the structural model (**d**). (**e-h**) The figure features the surroundings of the two chlorophylls of P680 and the accessory chlorophyll of the active electron transfer branch of PSII. The omit electron density map at the contour level of 1.5 σ for the dark (S_1) state is depicted in green (**e**) and the double flash (putative S_3) state is depicted in white (**f**). The figure also shows the overlay of the two omit maps (**g**) at the contour level of 1.5 σ. Model of the chlorophylls of the primary electron donor P680 without electron density map is also shown in (**h**).*

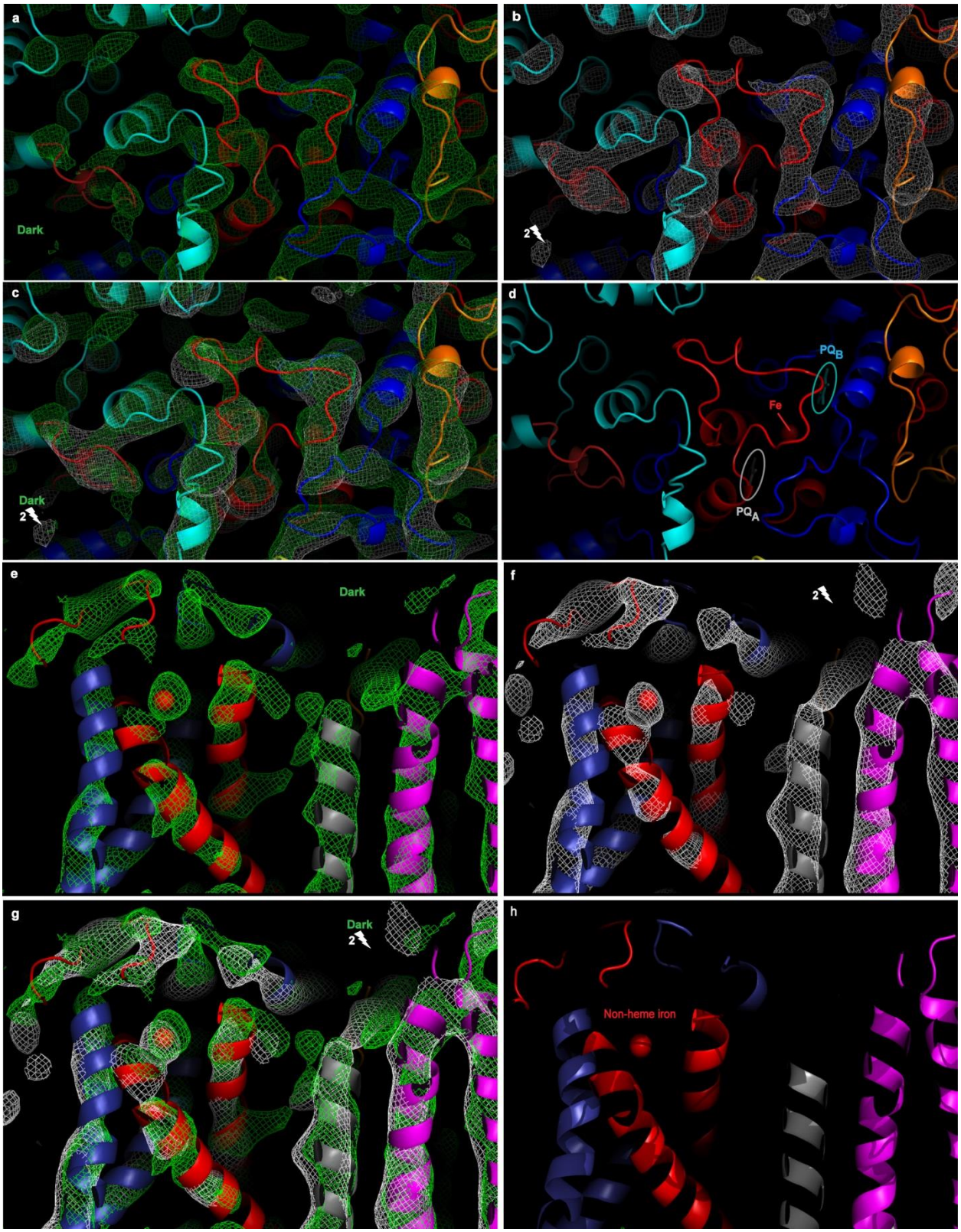

a
Dark
b
2
c
Dark
2
d
PQB
Fe
PQA
e
Dark
f
2
g
Dark
2
h
Non-heme iron

Figure A.3. The Electron Acceptor Side of PSII*. Omit map electron density and structural model of the dark and double-flash state of PSII – view from the stromal side onto the membrane plane. (**a-d**) These panels show the loops that coordinate the non-heme iron and cover the quinone binding sites looking from the stromal side onto the membrane plane. The omit electron density map at the contour level of 1.5 σ for the dark (S_1) state is depicted in green (**a**) and the double-flash (putative S_3) state is depicted in white (**b**). The figure also shows the overlay of the two omit maps (**c**) at 1.5 σ. The structural model **(d)** indicates the positions of PQ_A and PQ_B as well as the non-heme iron located below the loops. We note that the electron densities of the loop regions at the electron acceptor side show significant differences between the dark and the double flash states. The electron density of both states may suggest a conformation of the loops that could differ in their backbone trace from the model derived from the 1.9 Å structure from Umena et al.(Umena et al., 2011) (**e-h**) panels show the side view of the acceptor side of PSII along the plane in the membrane. The omit electron density map at the contour level of 1.5 σ for the dark (S_1) state is depicted in green (**e**) and the double-flash (putative S_3) state is depicted in white (**f**). The figure shows the overlay of the two omit maps (**g**) featuring the changes in the position of the non-heme iron and loop regions at the contour level of 1.5 σ. Model of the electron acceptor side of PSII (**h**). The protein subunits are color coded as in Fig. A.1; the non-heme iron is depicted as a red sphere. The tightly bound plastoquinone PQ_A is shown in white, the mobile plastoquinone PQ_B is depicted in cyan.*

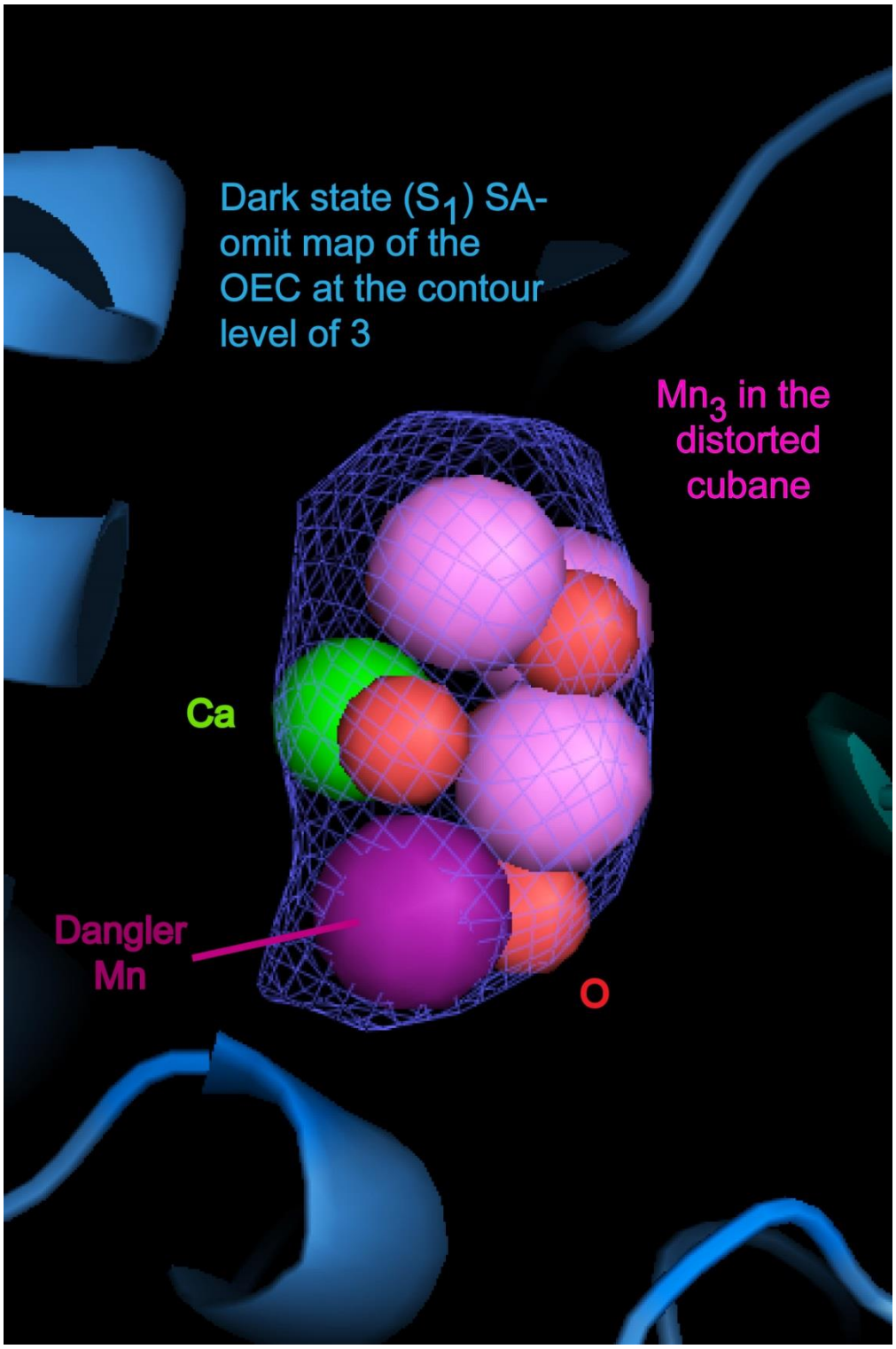


Figure A.4. Simulated Annealed Omit Map of the Mn_4CaO_5 Cluster of PSII. *The electron density of the dark state of PSII. This figure shows the superimposed SA-omit*

maps for the dark (S_1) (blue) state of the Mn_4CaO_5 cluster. It is noteworthy that we use a different color scheme for the SA-omit maps and the "regular $2F_o$-F_c" omits maps to allow the reader a better orientation of the type of map shown in each figure. The electron density is shown at the contour level of 3.0 σ to ensure that it solely features the metal Mn_4CaO_5 cluster. The X-ray structure at 1.9 Å from Umena et al.(Umena et al., 2011) is placed inside the SA-omit map for comparison. The nomenclature for the Mn atoms proposed in (Umena et al., 2011) is used for the color-coding of the individual Mn atoms in the cluster. The Mn ions that form the distorted Mn_3O_xCa cubane (Mn1, Mn2 and Mn3) are depicted in light pink, while Mn4 (violet) (named the "dangler" Mn) is located outside the cubane. This figure shows that the dangler Mn sticks out of the SA-omit map electron density, which indicates that this Mn atom may be located in closer proximity to the Mn_3CaO_X cubane in the dark S_1-state that is not influenced by X-ray damage.

APPENDIX C

PERMISSIONS

Reprinted by permission from Macmillan Publishers Ltd: [Nature] (**Neutze R, Wouts R, van der Spoel D, Weckert E, Hajdu J** (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. Nature **406:** 752-757), copyright (2000)

Reprinted by permission from Macmillan Publishers Ltd: [Nature Physics] (**Chapman HN, Barty A, Bogan MJ, Boutet S, Frank M, Hau-Riege S, Marchesini S, Woods BW, Bajt S, Benner WH, London RA, Plonjes E, Kuhlmann M, Treusch R, Dusterer S, Tschentscher T, Schneider JR, Spiller E, Moller T, Bostedt C, Hoener M, Shapiro D, Hodgson KO, Van der Spoel D, Burmeister F, Bergh M, Caleman C, Huldt G, Seibert M, Maia FR, Lee RW, Szoke A, Timneanu N, Hajdu J** (2006) Femtosecond Diffractive Imaging with a Soft-X-ray Free-Electron Laser. Nature Physics **2:** 839-843), copyright (2006)

Reprinted by permission from Macmillan Publishers Ltd: [Crystallography Reports] (**Altarelli M** (2010) From 3rd to 4th-Generation Light Sources: Free-Electron Lasers in the X-ray Range. Crystallography Reports **55:** 1145-1151), copyright (2010)

Reprinted by permission from Macmillan Publishers Ltd: [Nature] (**Umena Y, Kawakami K, Shen JR, Kamiya N** (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A. Nature **473:** 55-60), copyright (2011)

Reprinted by permission from Macmillan Publishers Ltd: [Nature] (**Loll B, Kern J, Saenger W, Zouni A, Biesiadka J** (2005) Towards complete cofactor arrangement in the 3.0 A resolution structure of photosystem II. Nature **438:** 1040-1044), copyright (2005)

Reprinted by permission from Macmillan Publishers Ltd: [Nature] (**Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N** (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution. Nature **411:** 909-917), copyright (2001)

Reprinted by permission from Macmillan Publishers Ltd: [Nature Protocols] (**Caffrey M, Cherezov V** (2009) Crystallizing membrane proteins using lipidic mesophases. Nat Protoc **4:** 706-731), copyright (2009)